Parameter Estimation in Diffusion Processes on the Space of Shapes

Valentina Staneva

joint work with Laurent Younes

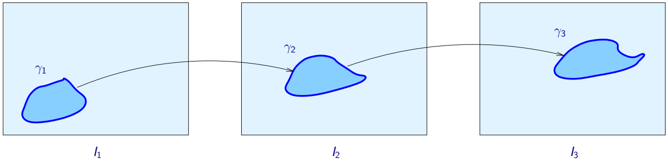
Applied Mathematics & Statistics Department
Center for Imaging Science
Johns Hopkins University[1]

September 4, 2014

[1] This research is partially supported by the ONR

Introduction

Our goal is to study temporal changes of shapes observed in videos.



We need to:

- construct stochastic processes for modeling the evolution of 2D shapes.
- provide methods for parameter estimation from a sequence of observations
- ensure the solutions are applicable to situations when observations are sparse

Background

- "The diffusion of Shape"(Kendall'77) introduces Brownian motion on the space of landmarks[1]
- extensions to an Ornstein - Uhlenbeck process (Ball'08)[2]
- modeling biological growth by diffeomorphisms (Grenander'06)[3]
- stochastic diffeomorphic flows for shapes (Vialard'12)[4]

The Geometry of Shape Space

- The *shape manifold* $\mathcal{M}$ consists of simple closed plane curves γ
  - here $\gamma = \{x_1, ..., x_m \in \mathbb{R}^2\}$.

The Geometry of Shape Space

- The *shape manifold* $\mathcal{M}$ consists of simple closed plane curves γ
  - here $\gamma = \{x_1, \ldots, x_m \in \mathbb{R}^2\}$.
- The *tangent space* $\mathcal{T}\mathcal{M}$ contains collections of 2D vectors
  - $\mathbf{v} = \{v_1, \ldots, v_m \in \mathbb{R}^2\}$

The Geometry of Shape Space

- The *shape manifold* $\mathcal{M}$ consists of simple closed plane curves γ
  - here $\gamma = \{x_1, ..., x_m \in \mathbb{R}^2\}$.
- The *tangent space* $\mathcal{TM}$ contains collections of 2D vectors
  - $\mathbf{v} = \{v_1, ..., v_m \in \mathbb{R}^2\}$
- *Riemannian metric:* $\|\mathbf{v}\|_\gamma = \mathbf{v}\mathbf{K}(\gamma)^{-1}\mathbf{v}$
  - $K(x_i, x_j) = \exp(\|x_i - x_j\|^2/2\sigma^2)$

The Geometry of Shape Space

- The *shape manifold* $\mathcal{M}$ consists of simple closed plane curves γ
  - here $\gamma = \{x_1, ..., x_m \in \mathbb{R}^2\}$.
- The *tangent space* $\mathcal{TM}$ contains collections of 2D vectors
  - $\mathbf{v} = \{v_1, ..., v_m \in \mathbb{R}^2\}$
- *Riemannian metric:* $\|\mathbf{v}\|_\gamma = \mathbf{v}\mathbf{K}(\gamma)^{-1}\mathbf{v}$
  - $K(x_i, x_j) = \exp(\|x_i - x_j\|^2/2\sigma^2)$
- *Exponential map:* $\gamma_1 = \exp_{\gamma_0}(\mathbf{v})$
  - can be computed by solving a system of ODE's

The Geometry of Shape Space

- The *shape manifold* $\mathcal{M}$ consists of simple closed plane curves γ
  - here $\gamma = \{x_1, ..., x_m \in \mathbb{R}^2\}$.
- The *tangent space* $\mathcal{TM}$ contains collections of 2D vectors
  - $\mathbf{v} = \{v_1, ..., v_m \in \mathbb{R}^2\}$
- *Riemannian metric:* $\|\mathbf{v}\|_\gamma = \mathbf{v}\mathbf{K}(\gamma)^{-1}\mathbf{v}$
  - $K(x_i, x_j) = \exp(\|x_i - x_j\|^2/2\sigma^2)$
- *Exponential map:* $\gamma_1 = \exp_{\gamma_0}(\mathbf{v})$
  - can be computed by solving a system of ODE's
- *Logarithm map:* $\mathbf{v} = \log_{\gamma_0}(\gamma_1)$
  - we assume it exists locally

The Geometry of Shape Space

- The *shape manifold* $\mathcal{M}$ consists of simple closed plane curves γ
  - here $\gamma = \{x_1, ..., x_m \in \mathbb{R}^2\}$.
- The *tangent space* $\mathcal{TM}$ contains collections of 2D vectors
  - $\mathbf{v} = \{v_1, ..., v_m \in \mathbb{R}^2\}$
- *Riemannian metric:* $\|\mathbf{v}\|_\gamma = \mathbf{v}\mathbf{K}(\gamma)^{-1}\mathbf{v}$
  - $K(x_i, x_j) = \exp(\|x_i - x_j\|^2/2\sigma^2)$
- *Exponential map:* $\gamma_1 = \exp_{\gamma_0}(\mathbf{v})$
  - can be computed by solving a system of ODE's
- *Logarithm map:* $\mathbf{v} = \log_{\gamma_0}(\gamma_1)$
  - we assume it exists locally

Note: in our setting we assume $v(\cdot) = \sum_{i=1}^{n} K(x_i, \cdot)\alpha_i$, where $\alpha_i \in \mathbb{R}^2$ and $\{x_i\}_{i=1}^{n}$ is a subset of the points on γ ($n < m$). This turns $\mathcal{M}$ into a *sub-Riemannian manifold*.

Diffusions on Manifolds

Let $\mathcal{M}$ be a m-dimensional manifold. Need to define

$$dX_t = A(X_t, \theta)dt + B(X_t)dW_t, \quad X_t \in \mathcal{M}.$$

- Itô and Stratonovich equations can be defined locally in a chart.
- Global formulation needs to take care of transformations between charts.

Diffusions on Manifolds

Let $\mathcal{M}$ be a m-dimensional manifold. Need to define

$$dX_t = A(X_t, \theta)dt + B(X_t)dW_t, \quad X_t \in \mathcal{M}.$$

- Itô and Stratonovich equations can be defined locally in a chart.
- Global formulation needs to take care of transformations between charts.

Let $\{A, B_1, ..., B_n\}$ be vector fields on $\mathcal{M}$.

Stratonovich Equations on $\mathcal{M}$

The solution is a process X_t on $\mathcal{M}$, which satisfies for any smooth compactly supported f

$$f(X_T) - f(X_0) = \int_0^T Af(X_t)dt + \int_0^T \sum_{k=1}^{n} B_k f(X_t) \circ dw_k(t)$$

Note: Stratonovich equations transform as tangent vectors!

Itô Equations on a Manifold

Itô equations transform according to Itô's Rule, not as tangent vectors.

Itô Equations on a Manifold

Itô equations transform according to Itô's Rule, not as tangent vectors.

$$f(X_T) - f(X_0) = \int_0^T Af(X_t)dt + \int_0^T \sum_{k=1}^{n} B_k f(X_t) \cdot dw_k(t)$$

Itô Equations on a Manifold

Itô equations transform according to Itô's Rule, not as tangent vectors.

$$\xcancel{f(X_T) - f(X_0) = \int_0^T Af(X_t)dt + \int_0^T \sum_{k=1}^{n} B_k f(X_t) \cdot dw_k(t)}$$

Instead we can define them through the exponential map.

Itô Equations in Belopolskaya-Daletsky form:

$$dX_t = \exp_{X_t}\left(A(X_t, \theta)dt + \sum_{k=1}^{n} B_k \cdot dw_k(t)\right),$$

Generating Diffusions

Consider the following vector fields on the shape manifold $K(\cdot, x_1)\mathbf{e}, \ldots, K(\cdot, x_n)\mathbf{e}$. We can obtain from them orthonormal vector fields: $E_1, \ldots, E_n$.

Generating Brownian motion on $\mathcal{M}$

$$X_{t+dt} = \exp_{X_t}\left(\sqrt{dt}\sum_{k=1}^{n} E_k(X_t)\varepsilon_k\right),$$

where $\{\varepsilon_k\}_{k=1}^{n} \sim \mathcal{N}(\mathbf{0}, \mathbf{I_n})$.

Constant Drift

Constant with respect to the **K** basis

$$dX_t = \sum_{k=1}^{n} \theta_k K(x_k, X_t)dt + dW_t$$

Note: the drift will not be constant wrt to other bases.

Mean-reverting Drift

Ornstein-Uhlenbeck process on $\mathbb{R}^n$

$$dX_t = -\theta \nabla_{X_t} \left[(X_t - \mu)^2 \right] + dW_t$$

Let μ be a template shape.

Mean-reverting process on $\mathcal{M}$

$$dX_t = -\theta \nabla_{X_t} dist(X_t, \mu) + dW_t$$

Instead of Riemannian distance we use:

$$dist(X_t, \mu) = \text{area of mismatch of the two shapes}$$

Length-area Drift

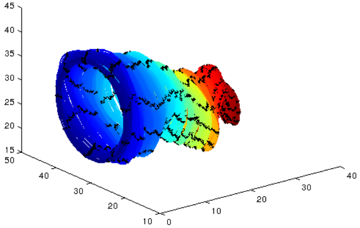
Suppose we don't have μ but only estimates for its length L_μ and area A_μ.

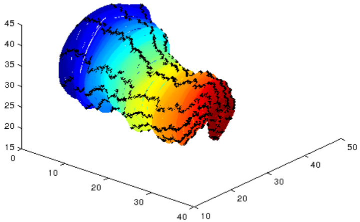
Shape process

$$dX_t = -\frac{1}{2}\theta_1 \nabla_{X_t} |L_{X_t} - L_\mu|^2 - \frac{1}{2}\theta_2 \nabla_{X_t} |A_{X_t} - A_\mu|^2 + dW_t$$

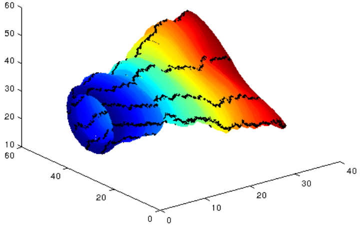
The above gradients can be explicitly computed using the notion of a shape gradient (Zolesio'81)[5].

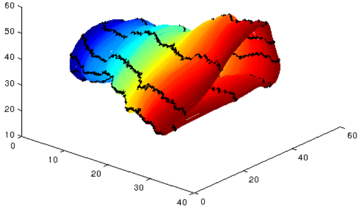
Driftless Diffusion

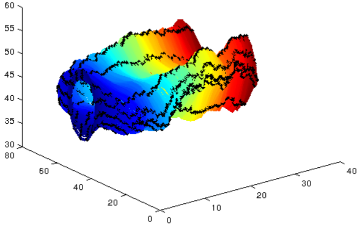




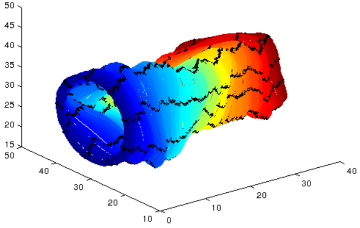
Constant Drift Diffusion

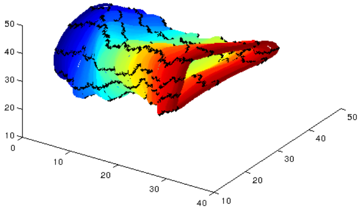





60
55
50
45
40
35
30
80
60
40
20
0
0
10
20
30
40
60
55
50
45
40
35
30
0
10
20
30
40
0
20
40
60
80

Length-Area Drift Diffusion





Parameter Estimation in Diffusions on $\mathbb{R}^n$

Let X_t be a diffusion process on $\mathbb{R}^n$:

$$\begin{aligned} \mu_X : \quad dX_t &= A(X_t, \theta)dt + dW_t \\ \mu_Y : \quad dY_t &= dW_t \end{aligned}$$

Girsanov Theorem:

Under some boundedness conditions on $A(X_t, \theta)$ we have $\mu_X \sim \mu_Y$ and

$$\frac{d\mu_X}{d\mu_Y}(X_t) = \exp\left(\int_0^T A(X_t, \theta)^T \cdot dX_t - \frac{1}{2}\int_0^T A(X_t, \theta)^T A(X_t, \theta) dt\right).$$

Parameter estimate of θ can be obtained by maximizing $\frac{d\mu_X}{d\mu_Y}(X_t)$ wrt θ.

Parameter Estimation in Diffusions on $\mathbb{R}^n$

Let X_t be a diffusion process on $\mathbb{R}^n$:

$$
\begin{aligned}
\mu_X : \quad dX_t &= A(X_t, \theta)dt + dW_t \\
\mu_Y : \quad dY_t &= dW_t
\end{aligned}
$$

Girsanov Theorem:

Under some boundedness conditions on $A(X_t, \theta)$ we have $\mu_X \sim \mu_Y$ and

$$\frac{d\mu_X}{d\mu_Y}(X_t) = \exp\left(\int_0^T A(X_t, \theta)^T \cdot dX_t - \frac{1}{2} \int_0^T A(X_t, \theta)^T A(X_t, \theta) dt \right).$$

Parameter estimate of θ can be obtained by maximizing $\frac{d\mu_X}{d\mu_Y}(X_t)$ wrt θ.

Note: when A is linear wrt θ, the exponent above becomes a quadratic function and there is a closed form solution for θ.

Girsanov Theorem on Manifolds (Elworthy'82)[6]

$$\frac{d\mu_X}{d\mu_Y}(X_t) = \exp\left(\int_0^T \langle A(X_t,\theta), \cdot dX_t\rangle_{X_t} - \frac{1}{2}\int_0^T \langle A(X_t,\theta), A(X_t,\theta)\rangle_{X_t} dt\right),$$

where $\cdot dX_t$ should be interpreted as $\log_{X_t}(X_{t+dt})$

Parameter Estimates

Given a sequence of observations $X_1, ..., X_N$ we can approximate the stochastic integrals.

- constant drift:

$$\hat{\theta} = \frac{1}{T}\sum_{i=1}^{N}(K_{X_i}^{-1/2})^T \log(X_i, X_{i+1})$$

- mean-reverting drift:

$$\hat{\theta} = \frac{\sum_{i=1}^{N-1}\langle \nabla dist(X_i,\mu), \log(X_i, X_{i+1})\rangle}{\sum_{i=1}^{N} \|\nabla dist(X_i,\mu)\|^2 dt}.$$

- length-area drift

$$\begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix} = \left(\sum_{i=1}^{N-1} M_i dt\right)^{-1} \begin{bmatrix} \sum_{i=1}^{N-1}\langle \nabla |L(X_i) - L_\mu|^2, \log(X_i, X_{i+1})\rangle \\ \sum_{i=1}^{N-1}\langle \nabla |A(X_i) - A_\mu|^2, \log(X_i, X_{i+1})\rangle \end{bmatrix},$$

where M_i as the Grammian matrix of $\nabla|L(X_i) - L|^2$ and $\nabla|A(X_i) - A|^2$.

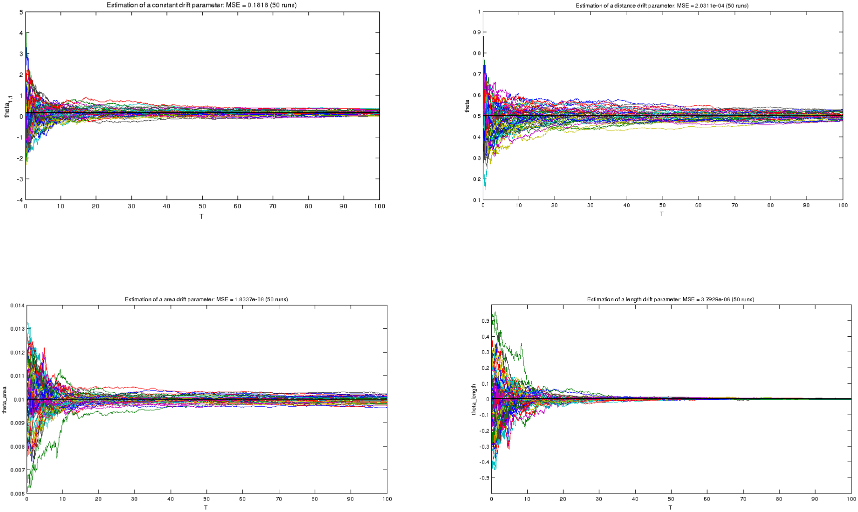


Figure: Top left: constant drift coefficients; top right: distance drift coefficient; bottom left: area drift coefficient; bottom right: length drift coefficient.

Current and Future Work

- parameter estimation from sparse observations
  - EM formulation to treat the missing observations
  - importance sampling approximation of the expectation step
- including higher order shape terms: e.g. curvature.
- considering the distance to several template shapes
- classification of time series

References

- D. Kendall, "The diffusion of shape," Advances in Applied Probability, vol. 9, no. 3, pp. 428–430, 1977.
- F. Ball, I. Dryden, and M. Golalizadeh, "Brownian motion and Ornstein-Uhlenbeck processes in planar shape space," Methodology and Computing in Applied Probability, vol. 10, no. 1, pp. 1–22, 2008.
- U. Grenander, A. Srivastava, and S. Saini, "Characterization of biological growth using iterated diffeomorphisms," IEEE International Symposium on Biological Imaging, pp. 1136–1139, 2006.
- F.-X. Vialard, "Extension to infinite dimensions of a stochastic second-order model associated with shape splines," Stochastic Processes and their Applications, vol. 123, no. 6, pp. 2110 – 2157, 2013.
- J. Zolesio, "The material derivative (or speed) method for shape optimization," Optimization of distributed parameter structures II, N.A.T.O. Asi Series, Series E: Applied Sciences, no. 50, pp. 1089–1151, 1981.
- K. Elworthy, Stochastic Differential Equations on Manifolds.