Differential Geometric Simulation Methods for Uncertainty Quantification in Complex Model Systems

Mark Girolami

Department of Statistics
University of Warwick

Warwick Centre for Predictive Modelling
February 2015

Never Mind the Big Data here's the Big Models

Never Mind the Big Data here's the Big Models

Never Mind the Big Data here's the Big Models

- Statistical inference in mathematical models (Finite Elements) emerging area of rapid development and growing importance in statistical methodology

Never Mind the Big Data here's the Big Models

- Statistical inference in mathematical models (Finite Elements) emerging area of rapid development and growing importance in statistical methodology
- Quantifying and propagating uncertainty in models e.g. Partial Differential Equations describing natural phenomena

Never Mind the Big Data here's the Big Models

- Statistical inference in mathematical models (Finite Elements) emerging area of rapid development and growing importance in statistical methodology
- Quantifying and propagating uncertainty in models e.g. Partial Differential Equations describing natural phenomena
- Examples include Shallow Water Components of Global Climate models (UK MET Office)

Never Mind the Big Data here's the Big Models

- Statistical inference in mathematical models (Finite Elements) emerging area of rapid development and growing importance in statistical methodology
- Quantifying and propagating uncertainty in models e.g. Partial Differential Equations describing natural phenomena
- Examples include Shallow Water Components of Global Climate models (UK MET Office)
- Ground water flow, subsurface multi-phase flow e.g. oil aquifers (BP, Schlumberger)

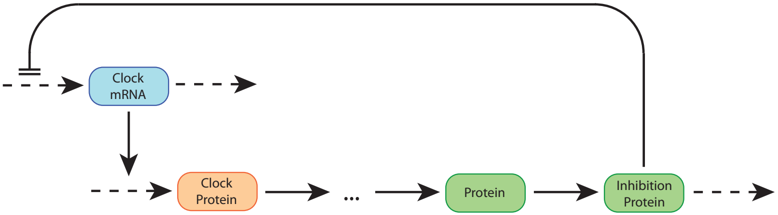
Never Mind the Big Data here's the Big Models

- Statistical inference in mathematical models (Finite Elements) emerging area of rapid development and growing importance in statistical methodology
- Quantifying and propagating uncertainty in models e.g. Partial Differential Equations describing natural phenomena
- Examples include Shallow Water Components of Global Climate models (UK MET Office)
- Ground water flow, subsurface multi-phase flow e.g. oil aquifers (BP, Schlumberger)
- Statistical inference in such models hugely challenging due to size of models and complexity of induced probability measures

Never Mind the Big Data here's the Big Models

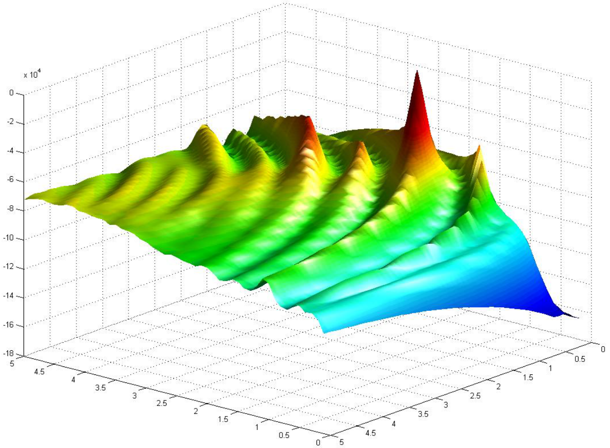
- Statistical inference in mathematical models (Finite Elements) emerging area of rapid development and growing importance in statistical methodology
- Quantifying and propagating uncertainty in models e.g. Partial Differential Equations describing natural phenomena
- Examples include Shallow Water Components of Global Climate models (UK MET Office)
- Ground water flow, subsurface multi-phase flow e.g. oil aquifers (BP, Schlumberger)
- Statistical inference in such models hugely challenging due to size of models and complexity of induced probability measures
- This talk will concentrate on the issues associated with sampling from the induced measures from such models and associated data

System Identification: Nonlinear ODE Oscillator Model

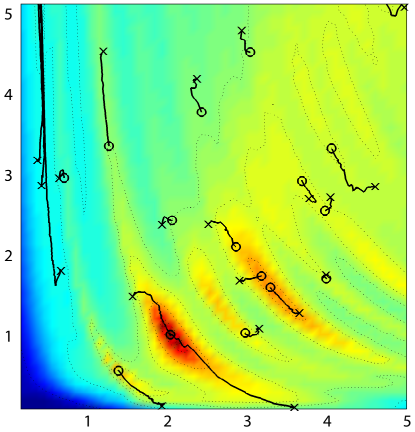


$$
\begin{aligned}
\frac{dx_1}{dt} &= \frac{k_1}{1 + x_n^{\rho}} - m_1 x_1 \\
\frac{dx_2}{dt} &= k_2 x_1 - m_2 x_2 \\
&\vdots \\
\frac{dx_n}{dt} &= k_n x_{n-1} - m_n x_n
\end{aligned}
$$

Systems Identification - Posterior Inference



Mixing of Markov Chains



Illustrative Heat Conduction Problem

- Heat conduction problem governed by an elliptic partial differential equation in the open and bounded domain Ω:

Illustrative Heat Conduction Problem

- Heat conduction problem governed by an elliptic partial differential equation in the open and bounded domain Ω:

$$\begin{aligned} -\nabla \cdot (e^u \nabla w) &= 0 \quad in \quad \Omega \\ -e^u \nabla w \cdot \mathrm{n} &= u \ \mathrm{Bi} \quad on \quad \partial\Omega / \Gamma_R \\ -e^u \nabla w \cdot \mathrm{n} &= -1 \quad on \quad \Gamma_R \end{aligned}$$

Illustrative Heat Conduction Problem

- Heat conduction problem governed by an elliptic partial differential equation in the open and bounded domain Ω:

$$\begin{aligned} -\nabla \cdot (e^{u} \nabla w) &= 0 \quad \mathit{in} \quad \Omega \\ -e^{u} \nabla w \cdot \mathrm{n} &= u \;\, \mathrm{Bi} \quad \mathit{on} \quad \partial\Omega / \, \Gamma_R \\ -e^{u} \nabla w \cdot \mathrm{n} &= -1 \quad \mathit{on} \quad \Gamma_R \end{aligned}$$

- Forward state is w and u the logarithm of distributed thermal conductivity on Ω, n the unit outward normal on Ω, and *Bi* the Biot number.

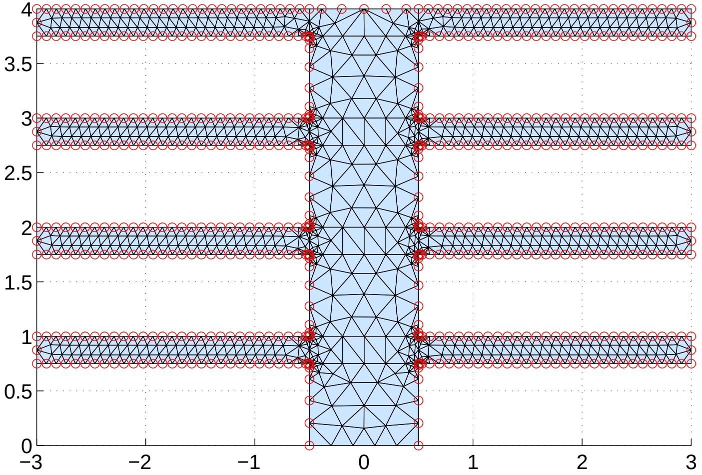
Illustrative Heat Conduction Problem

- Heat conduction problem governed by an elliptic partial differential equation in the open and bounded domain Ω:

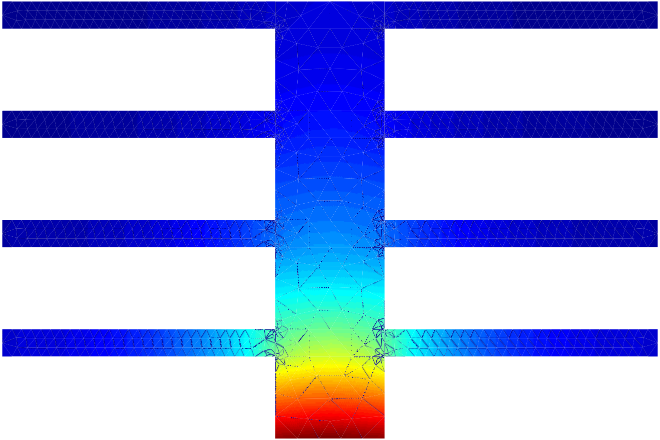
$$\begin{aligned} -\nabla \cdot (e^{u} \nabla w) &= 0 \quad \textit{in} \quad \Omega \\ -e^{u} \nabla w \cdot \mathrm{n} &= u \ \mathrm{Bi} \quad \textit{on} \quad \partial\Omega / \Gamma_R \\ -e^{u} \nabla w \cdot \mathrm{n} &= -1 \quad \textit{on} \quad \Gamma_R \end{aligned}$$

- Forward state is w and u the logarithm of distributed thermal conductivity on Ω, n the unit outward normal on Ω, and *Bi* the Biot number.

Illustrative Heat Conduction Problem



Illustrative Heat Conduction Problem

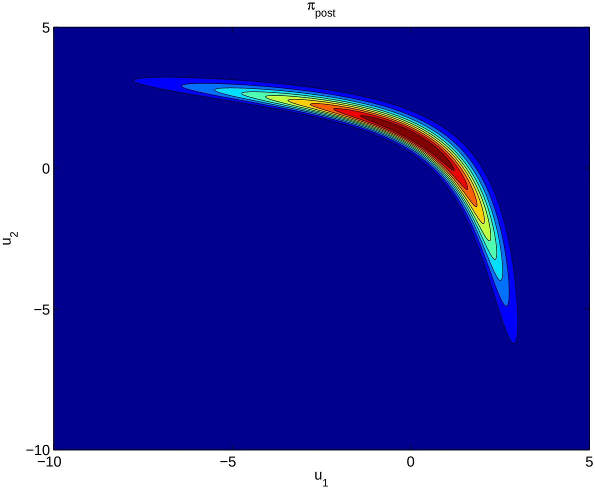


Illustrative Heat Conduction Problem

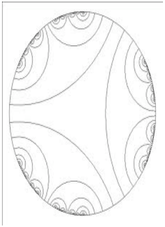
- Take one finite element discretisation of domain and one observation at leftmost boundary

Illustrative Heat Conduction Problem

- Take one finite element discretisation of domain and one observation at leftmost boundary
- Consider induced bivariate posterior for varying forms of prior Gaussian measure



MCMC from Diffusions and Geodesics


- Riemann manifold Langevin and Hamiltonian Monte Carlo Methods
  Girolami, M. & Calderhead, B.
  J.R.Statist. Soc. B, *with discussion*, (2011), **73**, *2*, 123 - 214.

`http://www2.warwick.ac.uk/mgirolami`

Talk Outline

- Motivation to improve MCMC capability for challenging problems

Talk Outline

- Motivation to improve MCMC capability for challenging problems
- Exploring differential geometric concepts in MCMC methodology

Talk Outline

- Motivation to improve MCMC capability for challenging problems
- Exploring differential geometric concepts in MCMC methodology
- Diffusions across Riemann manifold as basis for MCMC

Talk Outline

- Motivation to improve MCMC capability for challenging problems
- Exploring differential geometric concepts in MCMC methodology
- Diffusions across Riemann manifold as basis for MCMC
- Geodesic flows on manifold form basis of MCMC methods

Talk Outline

- Motivation to improve MCMC capability for challenging problems
- Exploring differential geometric concepts in MCMC methodology
- Diffusions across Riemann manifold as basis for MCMC
- Geodesic flows on manifold form basis of MCMC methods
- Illustrative examples

Talk Outline

- Motivation to improve MCMC capability for challenging problems
- Exploring differential geometric concepts in MCMC methodology
- Diffusions across Riemann manifold as basis for MCMC
- Geodesic flows on manifold form basis of MCMC methods
- Illustrative examples
- Further Work and Conclusions

Motivation Simulation Based Inference

- Monte Carlo method employs samples from $\pi(\boldsymbol{\theta})$ to obtain estimate

$$\int \phi(\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta} = \frac{1}{N}\sum\nolimits_n \phi(\boldsymbol{\theta}^n) + \mathcal{O}(N^{-\frac{1}{2}})$$

Motivation Simulation Based Inference

- Monte Carlo method employs samples from $\pi(\boldsymbol{\theta})$ to obtain estimate

$$\int \phi(\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta} = \frac{1}{N}\sum_n \phi(\boldsymbol{\theta}^n) + \mathcal{O}(N^{-\frac{1}{2}})$$

- Draw $\boldsymbol{\theta}^n$ from ergodic Markov process with stationary distribution $\pi(\boldsymbol{\theta})$

Motivation Simulation Based Inference

- Monte Carlo method employs samples from $\pi(\boldsymbol{\theta})$ to obtain estimate

$$\int \phi(\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta} = \frac{1}{N}\sum\nolimits_n \phi(\boldsymbol{\theta}^n) + \mathcal{O}(N^{-\frac{1}{2}})$$

- Draw $\boldsymbol{\theta}^n$ from ergodic Markov process with stationary distribution $\pi(\boldsymbol{\theta})$
- Construct transition kernel as product of two components

Motivation Simulation Based Inference

- Monte Carlo method employs samples from $\pi(\boldsymbol{\theta})$ to obtain estimate

$$\int \phi(\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta} = \frac{1}{N}\sum_n \phi(\boldsymbol{\theta}^n) + \mathcal{O}(N^{-\frac{1}{2}})$$

- Draw $\boldsymbol{\theta}^n$ from ergodic Markov process with stationary distribution $\pi(\boldsymbol{\theta})$
- Construct transition kernel as product of two components
  - Propose a move $\boldsymbol{\theta} \to \boldsymbol{\theta}'$ with probability $p_p(\boldsymbol{\theta}'|\boldsymbol{\theta})$

Motivation Simulation Based Inference

- Monte Carlo method employs samples from $\pi(\boldsymbol{\theta})$ to obtain estimate

$$\int \phi(\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta} = \frac{1}{N}\sum\nolimits_n \phi(\boldsymbol{\theta}^n) + \mathcal{O}(N^{-\frac{1}{2}})$$

- Draw $\boldsymbol{\theta}^n$ from ergodic Markov process with stationary distribution $\pi(\boldsymbol{\theta})$
- Construct transition kernel as product of two components
  - Propose a move $\boldsymbol{\theta} \to \boldsymbol{\theta}'$ with probability $p_p(\boldsymbol{\theta}'|\boldsymbol{\theta})$
  - accept or reject proposal with probability

$$p_a(\boldsymbol{\theta}'|\boldsymbol{\theta}) = \min\left[1, \frac{\pi(\boldsymbol{\theta}')p_p(\boldsymbol{\theta}|\boldsymbol{\theta}')}{\pi(\boldsymbol{\theta})p_p(\boldsymbol{\theta}'|\boldsymbol{\theta})}\right]$$

Motivation Simulation Based Inference

- Monte Carlo method employs samples from $\pi(\boldsymbol{\theta})$ to obtain estimate

$$\int \phi(\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta} = \frac{1}{N}\sum_n \phi(\boldsymbol{\theta}^n) + \mathcal{O}(N^{-\frac{1}{2}})$$

- Draw $\boldsymbol{\theta}^n$ from ergodic Markov process with stationary distribution $\pi(\boldsymbol{\theta})$
- Construct transition kernel as product of two components
  - Propose a move $\boldsymbol{\theta} \rightarrow \boldsymbol{\theta}'$ with probability $p_p(\boldsymbol{\theta}'|\boldsymbol{\theta})$
  - accept or reject proposal with probability

$$p_a(\boldsymbol{\theta}'|\boldsymbol{\theta}) = \min\left[1, \frac{\pi(\boldsymbol{\theta}')p_p(\boldsymbol{\theta}|\boldsymbol{\theta}')}{\pi(\boldsymbol{\theta})p_p(\boldsymbol{\theta}'|\boldsymbol{\theta})}\right]$$

- Convergence rate and asymptotic variance dependent on $p_p(\boldsymbol{\theta}'|\boldsymbol{\theta})$

Motivation Simulation Based Inference

- Monte Carlo method employs samples from $\pi(\boldsymbol{\theta})$ to obtain estimate

$$\int \phi(\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta} = \frac{1}{N}\sum\nolimits_n \phi(\boldsymbol{\theta}^n) + \mathcal{O}(N^{-\frac{1}{2}})$$

- Draw $\boldsymbol{\theta}^n$ from ergodic Markov process with stationary distribution $\pi(\boldsymbol{\theta})$
- Construct transition kernel as product of two components
  - Propose a move $\boldsymbol{\theta} \to \boldsymbol{\theta}'$ with probability $p_p(\boldsymbol{\theta}'|\boldsymbol{\theta})$
  - accept or reject proposal with probability

$$p_a(\boldsymbol{\theta}'|\boldsymbol{\theta}) = \min\left[1, \frac{\pi(\boldsymbol{\theta}')p_p(\boldsymbol{\theta}|\boldsymbol{\theta}')}{\pi(\boldsymbol{\theta})p_p(\boldsymbol{\theta}'|\boldsymbol{\theta})}\right]$$

- Convergence rate and asymptotic variance dependent on $p_p(\boldsymbol{\theta}'|\boldsymbol{\theta})$
- Success of MCMC reliant upon appropriate proposal design

Differential Geometric Concepts in MCMC

Differential Geometric Concepts in MCMC

- Denote expected Fisher Information as $\mathbf{G}(\boldsymbol{\theta}) = \text{cov}(\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}))$

Differential Geometric Concepts in MCMC

- Denote expected Fisher Information as $\mathbf{G}(\boldsymbol{\theta}) = \text{cov}(\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}))$
- Rao, 1945 to first order

$$\chi^2(\delta\boldsymbol{\theta}) = \int \frac{|p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) - p(\mathbf{y};\boldsymbol{\theta})|^2}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

Differential Geometric Concepts in MCMC

- Denote expected Fisher Information as $\mathbf{G}(\boldsymbol{\theta}) = \text{cov}(\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}))$
- Rao, 1945 to first order

$$\chi^2(\delta\boldsymbol{\theta}) = \int \frac{|p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) - p(\mathbf{y};\boldsymbol{\theta})|^2}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^\mathsf{T}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- Jeffreys, 1948 to first order

$$D(\boldsymbol{\theta}||\delta\boldsymbol{\theta}) = \int p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) \log \frac{p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta})}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^\mathsf{T}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

Differential Geometric Concepts in MCMC

- Denote expected Fisher Information as $\mathbf{G}(\boldsymbol{\theta}) = \text{cov}(\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}))$
- Rao, 1945 to first order

$$\chi^2(\delta\boldsymbol{\theta}) = \int \frac{|p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) - p(\mathbf{y};\boldsymbol{\theta})|^2}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- Jeffreys, 1948 to first order

$$D(\boldsymbol{\theta}||\delta\boldsymbol{\theta}) = \int p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) \log \frac{p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta})}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- Expected Fisher Information $\mathbf{G}(\boldsymbol{\theta})$ is metric tensor of a Riemann manifold

Differential Geometric Concepts in MCMC

- Denote expected Fisher Information as $\mathbf{G}(\boldsymbol{\theta}) = \text{cov}(\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}))$
- Rao, 1945 to first order

$$\chi^2(\delta\boldsymbol{\theta}) = \int \frac{|p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) - p(\mathbf{y};\boldsymbol{\theta})|^2}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- Jeffreys, 1948 to first order

$$D(\boldsymbol{\theta}||\delta\boldsymbol{\theta}) = \int p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) \log \frac{p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta})}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- Expected Fisher Information $\mathbf{G}(\boldsymbol{\theta})$ is metric tensor of a Riemann manifold
- Non-Euclidean geometry - invariants, connections, curvature, geodesics

Differential Geometric Concepts in MCMC

- Denote expected Fisher Information as $\mathbf{G}(\boldsymbol{\theta}) = \text{cov}(\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}))$
- Rao, 1945 to first order

$$\chi^2(\delta\boldsymbol{\theta}) = \int \frac{|p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) - p(\mathbf{y};\boldsymbol{\theta})|^2}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- Jeffreys, 1948 to first order

$$D(\boldsymbol{\theta}||\delta\boldsymbol{\theta}) = \int p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) \log \frac{p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta})}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- Expected Fisher Information $\mathbf{G}(\boldsymbol{\theta})$ is metric tensor of a Riemann manifold
- Non-Euclidean geometry - invariants, connections, curvature, geodesics
- Asymptotic statistical analysis. e.g. Amari, 1981; Murray & Rice, 1993; Critchley *et al*, 1993; Kass, 1989; Dawid, 1975; Lauritsen, 1989

Differential Geometric Concepts in MCMC

- Denote expected Fisher Information as $\mathbf{G}(\boldsymbol{\theta}) = \text{cov}(\nabla_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}))$
- Rao, 1945 to first order

$$\chi^2(\delta\boldsymbol{\theta}) = \int \frac{|p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) - p(\mathbf{y};\boldsymbol{\theta})|^2}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^\mathsf{T}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- Jeffreys, 1948 to first order

$$D(\boldsymbol{\theta}||\delta\boldsymbol{\theta}) = \int p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) \log \frac{p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta})}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^\mathsf{T}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- Expected Fisher Information $\mathbf{G}(\boldsymbol{\theta})$ is metric tensor of a Riemann manifold
- Non-Euclidean geometry - invariants, connections, curvature, geodesics
- Asymptotic statistical analysis. e.g. Amari, 1981; Murray & Rice, 1993; Critchley *et al*, 1993; Kass, 1989; Dawid, 1975; Lauritsen, 1989
- Statistical shape analysis Kent *et al*, 1996; Dryden & Mardia, 1998

Differential Geometric Concepts in MCMC

- Denote expected Fisher Information as $\mathbf{G}(\boldsymbol{\theta}) = \text{cov}(\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}))$
- Rao, 1945 to first order

$$\chi^2(\delta\boldsymbol{\theta}) = \int \frac{|p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) - p(\mathbf{y};\boldsymbol{\theta})|^2}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^\mathsf{T}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- Jeffreys, 1948 to first order

$$D(\boldsymbol{\theta}||\delta\boldsymbol{\theta}) = \int p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta}) \log \frac{p(\mathbf{y};\boldsymbol{\theta}+\delta\boldsymbol{\theta})}{p(\mathbf{y};\boldsymbol{\theta})} d\mathbf{y} \approx \delta\boldsymbol{\theta}^\mathsf{T}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

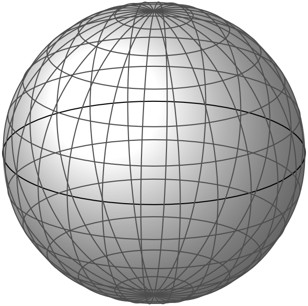
- Expected Fisher Information $\mathbf{G}(\boldsymbol{\theta})$ is metric tensor of a Riemann manifold
- Non-Euclidean geometry - invariants, connections, curvature, geodesics
- Asymptotic statistical analysis. e.g. Amari, 1981; Murray & Rice, 1993; Critchley *et al*, 1993; Kass, 1989; Dawid, 1975; Lauritsen, 1989
- Statistical shape analysis Kent *et al*, 1996; Dryden & Mardia, 1998

- Can geometric structure be employed in Monte Carlo methodology?

Manifolds

A manifold is a smooth, curved surface: A set *embedded* in $\mathbb{R}^d$, that locally looks like $\mathbb{R}^n$ ($n < d$).
Example: the unit sphere (2-sphere): $d = 3$, $n = 2$

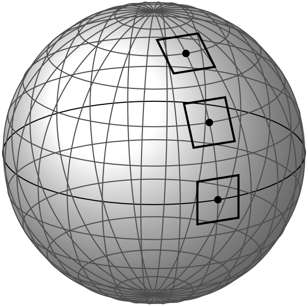
$$\mathcal{S}_2 = \{x \in \mathbb{R}^3 : \sum_i x_i^2 = 1\}$$



Manifolds

A manifold is a smooth, curved surface: A set *embedded* in $\mathbb{R}^d$, that locally looks like $\mathbb{R}^n$ ($n < d$).
Example: the unit sphere (2-sphere): $d = 3$, $n = 2$

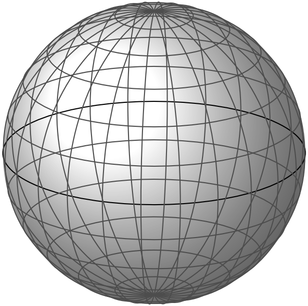
$$\mathcal{S}_2 = \{x \in \mathbb{R}^3 : \sum_i x_i^2 = 1\}$$

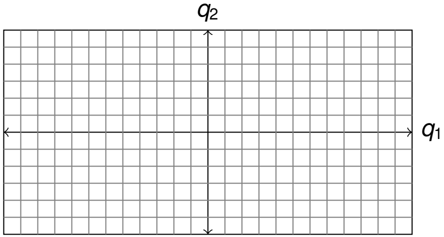


$x \in \mathbb{R}^d$ are the embedded coordinates

Coordinate systems and Riemannian metrics

Can parameterise the manifold with a coordinate system in $q \in \mathbb{R}^n$

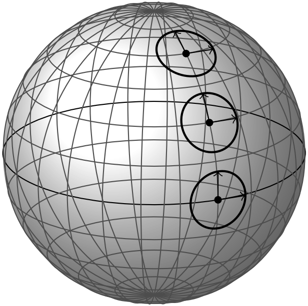


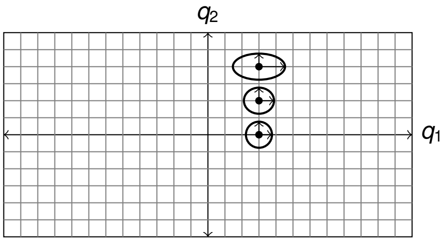


$$(\sin q_1 \sin q_2, \cos q_1 \sin q_2, \cos q_2), \qquad q_1 \in [0, 2\pi], q_2 \in [0, \pi]$$

Coordinate systems and Riemannian metrics

Can parameterise the manifold with a coordinate system in $q \in \mathbb{R}^n$



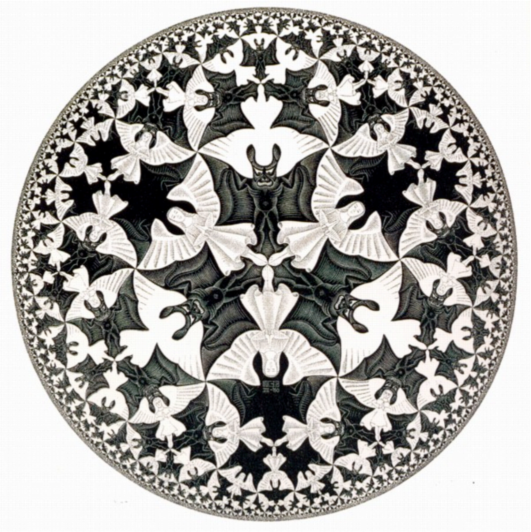


$$(\sin q_1 \sin q_2, \cos q_1 \sin q_2, \cos q_2), \qquad q_1 \in [0, 2\pi], q_2 \in [0, \pi]$$

The Euclidean metric $\|\cdot\|$ induces a Riemannian metric G in the coordinate system:

$$\|\mathrm{d}x\|^2 = \sum_{i,j} G(q)\mathrm{d}q_i\mathrm{d}q_j$$

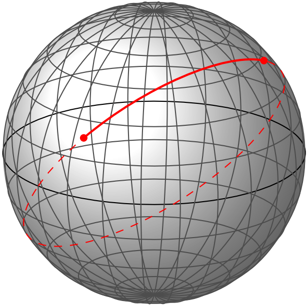
M.C. Escher, Heaven and Hell, 1960

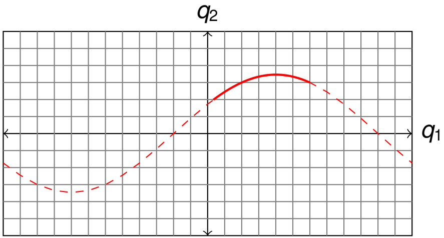


Geodesics

Geodesics are the paths of shortest distance.

- On a sphere, these are the *great circles*

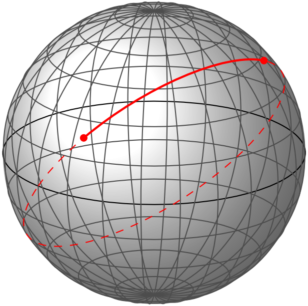


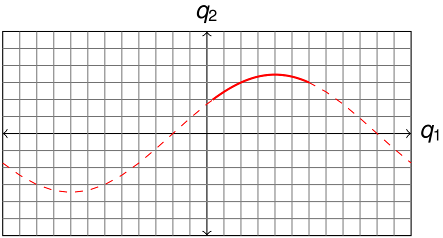


Geodesics

Geodesics are the paths of shortest distance.

- On a sphere, these are the *great circles*



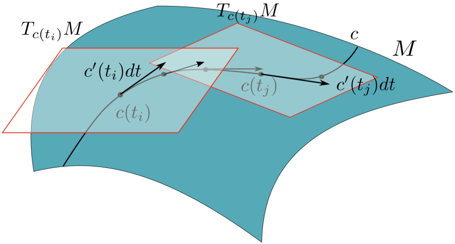


Given an initial velocity $v(0) \perp x(0)$, we have a nice explicit form

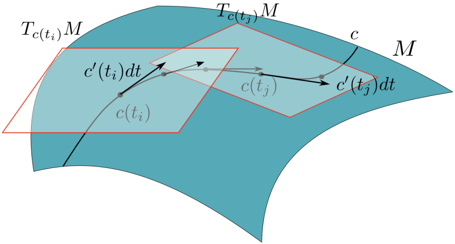
$$\begin{bmatrix} x(t) & v(t) \end{bmatrix} = \begin{bmatrix} x(0) & v(0) \end{bmatrix} \begin{bmatrix} 1 & \\ & \alpha^{-1} \end{bmatrix} \begin{bmatrix} \cos(\alpha t) & -\sin(\alpha t) \\ \sin(\alpha t) & \cos(\alpha t) \end{bmatrix} \begin{bmatrix} 1 & \\ & \alpha \end{bmatrix}$$

where $\alpha = \|v(0)\|$.

Geometric Concepts in MCMC

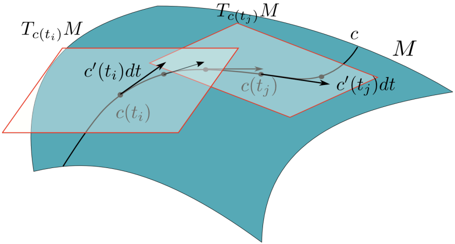


Geometric Concepts in MCMC



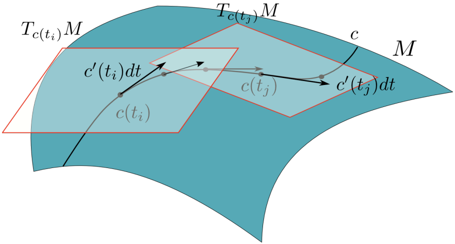
- Tangent space - local metric defined by $\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta} = g_{kl}\delta\theta_k\delta\theta_l$

Geometric Concepts in MCMC



- Tangent space - local metric defined by $\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta} = g_{kl}\delta\theta_k\delta\theta_l$
- Christoffel symbols - characterise Levi-Civita connection on manifold

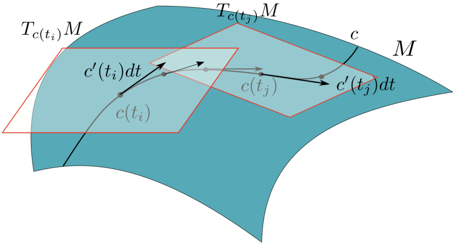
Geometric Concepts in MCMC



- Tangent space - local metric defined by $\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta} = g_{kl}\delta\theta_k\delta\theta_l$
- Christoffel symbols - characterise Levi-Civita connection on manifold

$$\Gamma^i_{kl} = \frac{1}{2}\sum_m g^{im}\left(\frac{\partial g_{mk}}{\partial\theta^l} + \frac{\partial g_{ml}}{\partial\theta^k} - \frac{\partial g_{kl}}{\partial\theta^m}\right)$$

Geometric Concepts in MCMC

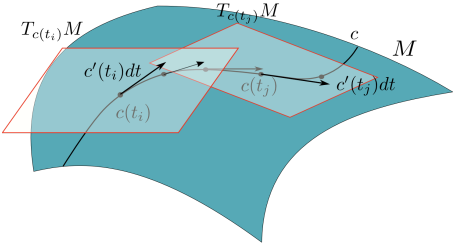


- Tangent space - local metric defined by $\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta} = g_{kl}\delta\theta_k\delta\theta_l$
- Christoffel symbols - characterise Levi-Civita connection on manifold

$$\Gamma^i_{kl} = \frac{1}{2}\sum_m g^{im}\left(\frac{\partial g_{mk}}{\partial\theta^l} + \frac{\partial g_{ml}}{\partial\theta^k} - \frac{\partial g_{kl}}{\partial\theta^m}\right)$$

- Geodesics - shortest path between two points on manifold

Geometric Concepts in MCMC



- Tangent space - local metric defined by $\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta} = g_{kl}\delta\theta_k\delta\theta_l$
- Christoffel symbols - characterise Levi-Civita connection on manifold

$$\Gamma^i_{kl} = \frac{1}{2}\sum_m g^{im}\left(\frac{\partial g_{mk}}{\partial\theta^l} + \frac{\partial g_{ml}}{\partial\theta^k} - \frac{\partial g_{kl}}{\partial\theta^m}\right)$$

- Geodesics - shortest path between two points on manifold

$$\frac{d^2\theta^i}{dt^2} + \sum_{k,l}\Gamma^i_{kl}\frac{d\theta^k}{dt}\frac{d\theta^l}{dt} = 0$$

Fisher–Rao metric

A family of probability densities $\{p(\cdot \mid \theta) : \theta \in \Theta\}$, the Fisher information forms a natural Riemannian metric, known as Fisher–Rao metric:

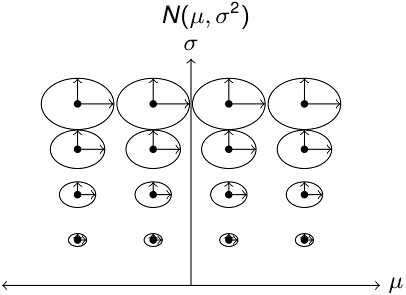
$$g_{ij} = E_{Y|\theta}\left[-\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log p(Y \mid \theta)\right]$$

Fisher–Rao metric

A family of probability densities $\{p(\cdot \mid \theta) : \theta \in \Theta\}$, the Fisher information forms a natural Riemannian metric, known as Fisher–Rao metric:

$$g_{ij} = E_{Y|\theta}\left[-\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log p(Y \mid \theta)\right]$$

Examples:

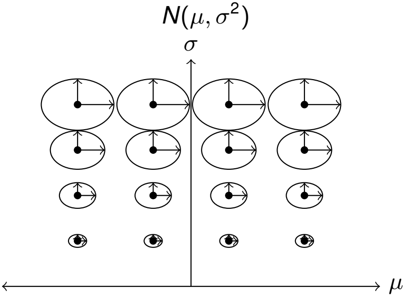


Fisher–Rao metric

A family of probability densities $\{p(\cdot \mid \theta) : \theta \in \Theta\}$, the Fisher information forms a natural Riemannian metric, known as Fisher–Rao metric:

$$g_{ij} = E_{Y|\theta}\left[-\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log p(Y \mid \theta)\right]$$

Examples:



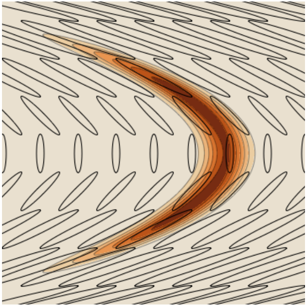


Illustration of Geometric Concepts

- Consider Normal density $p(x|\mu, \sigma) = \mathcal{N}_x(\mu, \sigma)$

Illustration of Geometric Concepts

- Consider Normal density $p(x|\mu, \sigma) = \mathcal{N}_x(\mu, \sigma)$
- Local inner product on tangent space defined by metric tensor, i.e. $\delta\boldsymbol{\theta}^\mathsf{T}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$, where $\boldsymbol{\theta} = (\mu, \sigma)^\mathsf{T}$

Illustration of Geometric Concepts

- Consider Normal density $p(x|\mu, \sigma) = \mathcal{N}_x(\mu, \sigma)$
- Local inner product on tangent space defined by metric tensor, i.e. $\delta\boldsymbol{\theta}^\mathsf{T}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$, where $\boldsymbol{\theta} = (\mu, \sigma)^\mathsf{T}$
- Metric is Expected Fisher Information

$$\mathbf{G}(\mu, \sigma) = \begin{bmatrix} \sigma^{-2} & 0 \\ 0 & 2\sigma^{-2} \end{bmatrix}$$

Illustration of Geometric Concepts

- Consider Normal density $p(x|\mu, \sigma) = \mathcal{N}_x(\mu, \sigma)$
- Local inner product on tangent space defined by metric tensor, i.e. $\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$, where $\boldsymbol{\theta} = (\mu, \sigma)^{\mathsf{T}}$
- Metric is Expected Fisher Information

$$\mathbf{G}(\mu, \sigma) = \begin{bmatrix} \sigma^{-2} & 0 \\ 0 & 2\sigma^{-2} \end{bmatrix}$$

- Components of connection $\partial_\mu \mathbf{G} = \mathbf{0}$ and $\partial_\sigma \mathbf{G} = -\operatorname{diag}(2\sigma^{-3}, 4\sigma^{-3})$

Illustration of Geometric Concepts

- Consider Normal density $p(x|\mu, \sigma) = \mathcal{N}_x(\mu, \sigma)$
- Local inner product on tangent space defined by metric tensor, i.e. $\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$, where $\boldsymbol{\theta} = (\mu, \sigma)^{\mathsf{T}}$
- Metric is Expected Fisher Information

$$\mathbf{G}(\mu, \sigma) = \begin{bmatrix} \sigma^{-2} & 0 \\ 0 & 2\sigma^{-2} \end{bmatrix}$$

- Components of connection $\partial_\mu \mathbf{G} = \mathbf{0}$ and $\partial_\sigma \mathbf{G} = -\,\mathrm{diag}(2\sigma^{-3}, 4\sigma^{-3})$
- Metric on tangent space

$$\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta} = \frac{(\delta\mu^2 + 2\delta\sigma^2)}{\sigma^2}$$

Illustration of Geometric Concepts

- Consider Normal density $p(x|\mu, \sigma) = \mathcal{N}_x(\mu, \sigma)$
- Local inner product on tangent space defined by metric tensor, i.e. $\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$, where $\boldsymbol{\theta} = (\mu, \sigma)^{\mathsf{T}}$
- Metric is Expected Fisher Information

$$\mathbf{G}(\mu, \sigma) = \begin{bmatrix} \sigma^{-2} & 0 \\ 0 & 2\sigma^{-2} \end{bmatrix}$$

- Components of connection $\partial_\mu \mathbf{G} = \mathbf{0}$ and $\partial_\sigma \mathbf{G} = -\operatorname{diag}(2\sigma^{-3}, 4\sigma^{-3})$
- Metric on tangent space

$$\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta} = \frac{(\delta\mu^2 + 2\delta\sigma^2)}{\sigma^2}$$

- Metric tensor for univariate Normal defines a Hyperbolic Space

Illustration of Geometric Concepts

- Consider Normal density $p(x|\mu, \sigma) = \mathcal{N}_x(\mu, \sigma)$
- Local inner product on tangent space defined by metric tensor, i.e. $\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta}$, where $\boldsymbol{\theta} = (\mu, \sigma)^{\mathsf{T}}$
- Metric is Expected Fisher Information

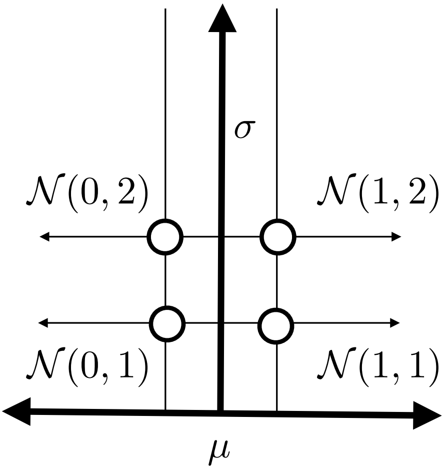
$$\mathbf{G}(\mu, \sigma) = \begin{bmatrix} \sigma^{-2} & 0 \\ 0 & 2\sigma^{-2} \end{bmatrix}$$

- Components of connection $\partial_\mu \mathbf{G} = \mathbf{0}$ and $\partial_\sigma \mathbf{G} = -\,\mathrm{diag}(2\sigma^{-3}, 4\sigma^{-3})$
- Metric on tangent space

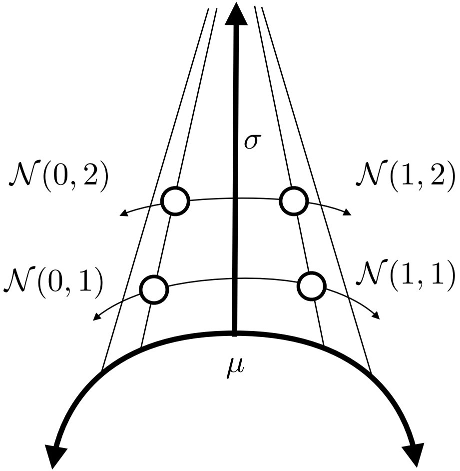
$$\delta\boldsymbol{\theta}^{\mathsf{T}}\mathbf{G}(\boldsymbol{\theta})\delta\boldsymbol{\theta} = \frac{(\delta\mu^2 + 2\delta\sigma^2)}{\sigma^2}$$

- Metric tensor for univariate Normal defines a Hyperbolic Space
- Consider densities $\mathcal{N}(0, 1)$ & $\mathcal{N}(1, 1)$ and $\mathcal{N}(0, 2)$ & $\mathcal{N}(1, 2)$

Normal Density - Euclidean Parameter space



Normal Density - Riemannian Functional space



Langevin Diffusion on Riemannian manifold

- Continuous Langevin diffusion with invariant measure $\pi(\boldsymbol{\theta}) \equiv \exp(\mathcal{L}(\boldsymbol{\theta}))$

$$d\boldsymbol{\theta} = \frac{1}{2}\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}) - \sum_{i,j}^{D} \boldsymbol{G}(\boldsymbol{\theta})^{-1}_{ij}\Gamma_{ij} + \sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}d\mathbf{W}$$

Langevin Diffusion on Riemannian manifold

- Continuous Langevin diffusion with invariant measure $\pi(\boldsymbol{\theta}) \equiv \exp(\mathcal{L}(\boldsymbol{\theta}))$

$$d\boldsymbol{\theta} = \frac{1}{2}\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}) - \sum_{i,j}^{D} \boldsymbol{G}(\boldsymbol{\theta})_{ij}^{-1}\Gamma_{ij} + \sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}d\mathbf{W}$$

- Discretised Langevin diffusion on manifold defines proposal mechanism

$$\boldsymbol{\theta}_d' = \boldsymbol{\theta}_d + \frac{\epsilon^2}{2}\left(\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta})\right)_d - \epsilon^2 \sum_{i,j}^{D} \boldsymbol{G}(\boldsymbol{\theta})_{ij}^{-1}\Gamma_{ij}^{d} + \epsilon\left(\sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}\mathbf{z}\right)_d$$

Langevin Diffusion on Riemannian manifold

- Continuous Langevin diffusion with invariant measure $\pi(\boldsymbol{\theta}) \equiv \exp(\mathcal{L}(\boldsymbol{\theta}))$

$$d\boldsymbol{\theta} = \frac{1}{2}\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}) - \sum_{i,j}^{D} \boldsymbol{G}(\boldsymbol{\theta})_{ij}^{-1}\Gamma_{ij} + \sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}d\mathbf{W}$$

- Discretised Langevin diffusion on manifold defines proposal mechanism

$$\boldsymbol{\theta}'_d = \boldsymbol{\theta}_d + \frac{\epsilon^2}{2}\left(\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta})\right)_d - \epsilon^2 \sum_{i,j}^{D} \boldsymbol{G}(\boldsymbol{\theta})_{ij}^{-1}\Gamma_{ij}^{d} + \epsilon\left(\sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}\mathbf{z}\right)_d$$

- Manifold with constant curvature then proposal mechanism reduces to

$$\boldsymbol{\theta}' = \boldsymbol{\theta} + \frac{\epsilon^2}{2}\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}) + \epsilon\sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}\mathbf{z}$$

Langevin Diffusion on Riemannian manifold

- Continuous Langevin diffusion with invariant measure $\pi(\boldsymbol{\theta}) \equiv \exp(\mathcal{L}(\boldsymbol{\theta}))$

$$d\boldsymbol{\theta} = \frac{1}{2}\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}) - \sum_{i,j}^{D}\boldsymbol{G}(\boldsymbol{\theta})_{ij}^{-1}\Gamma_{ij} + \sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}d\mathbf{W}$$

- Discretised Langevin diffusion on manifold defines proposal mechanism

$$\boldsymbol{\theta}_d' = \boldsymbol{\theta}_d + \frac{\epsilon^2}{2}\left(\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta})\right)_d - \epsilon^2\sum_{i,j}^{D}\boldsymbol{G}(\boldsymbol{\theta})_{ij}^{-1}\Gamma_{ij}^{d} + \epsilon\left(\sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}\mathbf{z}\right)_d$$

- Manifold with constant curvature then proposal mechanism reduces to

$$\boldsymbol{\theta}' = \boldsymbol{\theta} + \frac{\epsilon^2}{2}\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}) + \epsilon\sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}\mathbf{z}$$

- MALA proposal with preconditioning

$$\boldsymbol{\theta}' = \boldsymbol{\theta} + \frac{\epsilon^2}{2}\mathbf{M}\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}) + \epsilon\sqrt{\mathbf{M}}\mathbf{z}$$

Langevin Diffusion on Riemannian manifold

- Continuous Langevin diffusion with invariant measure $\pi(\boldsymbol{\theta}) \equiv \exp(\mathcal{L}(\boldsymbol{\theta}))$

$$d\boldsymbol{\theta} = \frac{1}{2}\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}) - \sum_{i,j}^{D} \boldsymbol{G}(\boldsymbol{\theta})^{-1}_{ij}\Gamma_{ij} + \sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}d\mathbf{W}$$

- Discretised Langevin diffusion on manifold defines proposal mechanism

$$\boldsymbol{\theta}'_d = \boldsymbol{\theta}_d + \frac{\epsilon^2}{2}\left(\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta})\right)_d - \epsilon^2 \sum_{i,j}^{D} \boldsymbol{G}(\boldsymbol{\theta})^{-1}_{ij}\Gamma^d_{ij} + \epsilon\left(\sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}\mathbf{z}\right)_d$$

- Manifold with constant curvature then proposal mechanism reduces to

$$\boldsymbol{\theta}' = \boldsymbol{\theta} + \frac{\epsilon^2}{2}\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}) + \epsilon\sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}\mathbf{z}$$

- MALA proposal with preconditioning

$$\boldsymbol{\theta}' = \boldsymbol{\theta} + \frac{\epsilon^2}{2}\mathbf{M}\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}) + \epsilon\sqrt{\mathbf{M}}\mathbf{z}$$

- Proposal and acceptance probability

$$p_p(\boldsymbol{\theta}'|\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\theta}'|\boldsymbol{\mu}(\boldsymbol{\theta}, \epsilon), \epsilon^2\mathbf{G}^{-1}(\boldsymbol{\theta}))$$

$$p_a(\boldsymbol{\theta}'|\boldsymbol{\theta}) = \min\left[1, \frac{\pi(\boldsymbol{\theta}')p_p(\boldsymbol{\theta}|\boldsymbol{\theta}')}{\pi(\boldsymbol{\theta})p_p(\boldsymbol{\theta}'|\boldsymbol{\theta})}\right]$$

Langevin Diffusion on Riemannian manifold

- Continuous Langevin diffusion with invariant measure $\pi(\boldsymbol{\theta}) \equiv \exp(\mathcal{L}(\boldsymbol{\theta}))$

$$d\boldsymbol{\theta} = \frac{1}{2}\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}) - \sum_{i,j}^{D}\boldsymbol{G}(\boldsymbol{\theta})_{ij}^{-1}\Gamma_{ij} + \sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}d\mathbf{W}$$

- Discretised Langevin diffusion on manifold defines proposal mechanism

$$\boldsymbol{\theta}'_d = \boldsymbol{\theta}_d + \frac{\epsilon^2}{2}\left(\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta})\right)_d - \epsilon^2\sum_{i,j}^{D}\boldsymbol{G}(\boldsymbol{\theta})_{ij}^{-1}\Gamma_{ij}^{d} + \epsilon\left(\sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}\mathbf{z}\right)_d$$

- Manifold with constant curvature then proposal mechanism reduces to

$$\boldsymbol{\theta}' = \boldsymbol{\theta} + \frac{\epsilon^2}{2}\boldsymbol{G}^{-1}(\boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}) + \epsilon\sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta})}\mathbf{z}$$

- MALA proposal with preconditioning

$$\boldsymbol{\theta}' = \boldsymbol{\theta} + \frac{\epsilon^2}{2}\mathbf{M}\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}) + \epsilon\sqrt{\mathbf{M}}\mathbf{z}$$

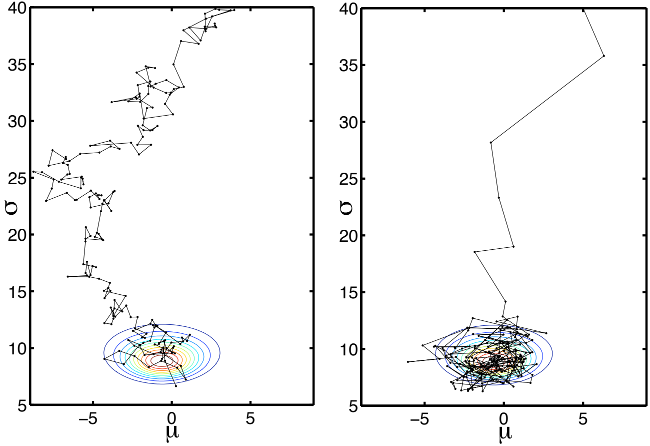
- Proposal and acceptance probability

$$p_p(\boldsymbol{\theta}'|\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\theta}'|\boldsymbol{\mu}(\boldsymbol{\theta},\epsilon), \epsilon^2\mathbf{G}^{-1}(\boldsymbol{\theta}))$$

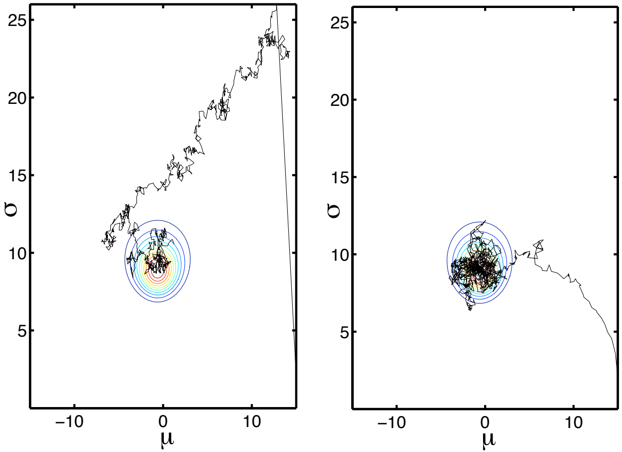
$$p_a(\boldsymbol{\theta}'|\boldsymbol{\theta}) = \min\left[1, \frac{\pi(\boldsymbol{\theta}')p_p(\boldsymbol{\theta}|\boldsymbol{\theta}')}{\pi(\boldsymbol{\theta})p_p(\boldsymbol{\theta}'|\boldsymbol{\theta})}\right]$$

- Proposal mechanism diffuses approximately along the manifold

Langevin Diffusion on Riemannian manifold



Langevin Diffusion on Riemannian manifold



Second-Order Diffusions

Second-Order Diffusions

- Define a second-order SDE on a Riemann manifold with the diffusion defined on vector field $\mathbf{v}$ to obtain

Second-Order Diffusions

- Define a second-order SDE on a Riemann manifold with the diffusion defined on vector field **v** to obtain

$$\begin{aligned} d\mathbf{x}_t &= \mathbf{v}_t dt \\ d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t \end{aligned}$$

Second-Order Diffusions

- Define a second-order SDE on a Riemann manifold with the diffusion defined on vector field **v** to obtain

$$\begin{aligned} d\mathbf{x}_t &= \mathbf{v}_t dt \\ d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t \end{aligned}$$

- Vector function γ has components $\gamma_k = \sum_{i,j} \Gamma_{ij}^k v^i v^j$

Second-Order Diffusions

- Define a second-order SDE on a Riemann manifold with the diffusion defined on vector field $\mathbf{v}$ to obtain

$$\begin{aligned} d\mathbf{x}_t &= \mathbf{v}_t dt \\ d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t \end{aligned}$$

- Vector function γ has components $\gamma_k = \sum_{i,j} \Gamma_{ij}^k v^i v^j$
- Each $\Gamma_{i,j}^k$ being the Christofell symbols of the Levi-Civita connection corresponding to the metric tensor $\mathbf{G}(\mathbf{x}_t)$

Second-Order Diffusions

- Define a second-order SDE on a Riemann manifold with the diffusion defined on vector field $\mathbf{v}$ to obtain

$$\begin{aligned} d\mathbf{x}_t &= \mathbf{v}_t dt \\ d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t \end{aligned}$$

- Vector function γ has components $\gamma_k = \sum_{i,j} \Gamma_{ij}^k v^i v^j$
- Each $\Gamma_{i,j}^k$ being the Christofell symbols of the Levi-Civita connection corresponding to the metric tensor $\mathbf{G}(\mathbf{x}_t)$
- Note $d\mathbf{v}_t + \gamma(\mathbf{x}_t, \mathbf{v}_t)dt$ is the covariant time derivative of velocity and the contravariant form of the potential gradient is simply $\mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)$

Second-Order Diffusions

- Define a second-order SDE on a Riemann manifold with the diffusion defined on vector field $\mathbf{v}$ to obtain

$$\begin{aligned} d\mathbf{x}_t &= \mathbf{v}_t dt \\ d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t \end{aligned}$$

- Vector function γ has components $\gamma_k = \sum_{i,j} \Gamma_{ij}^k v^i v^j$
- Each $\Gamma_{i,j}^k$ being the Christofell symbols of the Levi-Civita connection corresponding to the metric tensor $\mathbf{G}(\mathbf{x}_t)$
- Note $d\mathbf{v}_t + \gamma(\mathbf{x}_t, \mathbf{v}_t)dt$ is the covariant time derivative of velocity and the contravariant form of the potential gradient is simply $\mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)$
- Physically defines acceleration of a particle across manifold under the influence of potential field

Second-Order Diffusions

- Define a second-order SDE on a Riemann manifold with the diffusion defined on vector field $\mathbf{v}$ to obtain

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{v}_t dt \\
d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- Vector function γ has components $\gamma_k = \sum_{i,j} \Gamma_{ij}^k v^i v^j$
- Each $\Gamma_{i,j}^k$ being the Christofell symbols of the Levi-Civita connection corresponding to the metric tensor $\mathbf{G}(\mathbf{x}_t)$
- Note $d\mathbf{v}_t + \gamma(\mathbf{x}_t, \mathbf{v}_t)dt$ is the covariant time derivative of velocity and the contravariant form of the potential gradient is simply $\mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)$
- Physically defines acceleration of a particle across manifold under the influence of potential field
- A unit friction or viscous force is included by the term $\mathbf{v}$ and the fluctuating forcing is modelled by the Brownian motion

Second-Order Diffusions

- Define a second-order SDE on a Riemann manifold with the diffusion defined on vector field $\mathbf{v}$ to obtain

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{v}_t dt \\
d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- Vector function γ has components $\gamma_k = \sum_{i,j} \Gamma_{ij}^k v^i v^j$
- Each $\Gamma_{i,j}^k$ being the Christofell symbols of the Levi-Civita connection corresponding to the metric tensor $\mathbf{G}(\mathbf{x}_t)$
- Note $d\mathbf{v}_t + \gamma(\mathbf{x}_t, \mathbf{v}_t)dt$ is the covariant time derivative of velocity and the contravariant form of the potential gradient is simply $\mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)$
- Physically defines acceleration of a particle across manifold under the influence of potential field
- A unit friction or viscous force is included by the term $\mathbf{v}$ and the fluctuating forcing is modelled by the Brownian motion
- What has this to do with MCMC?

Second-Order Diffusions

Second-Order Diffusions

- What is invariant measure of this diffusion?

Second-Order Diffusions

- What is invariant measure of this diffusion?
- The Fokker-Planck equation for the diffusion follows as

Second-Order Diffusions

- What is invariant measure of this diffusion?
- The Fokker-Planck equation for the diffusion follows as

$$\frac{dp}{dt} = - v^a \frac{\partial p}{\partial x^a} + \frac{\partial}{\partial v^a}\left[\left(\Gamma^a_{bc} v^b v^c + g^{ab}\frac{\partial \phi}{\partial x^b} + v^a\right) p\right] + g^{ab}\frac{\partial^2 p}{\partial v^a \partial v^b}$$

Second-Order Diffusions

- What is invariant measure of this diffusion?
- The Fokker-Planck equation for the diffusion follows as

$$\frac{dp}{dt} = \quad - \quad v^a \frac{\partial p}{\partial x^a} + \frac{\partial}{\partial v^a}\left[\left(\Gamma^a_{bc} v^b v^c + g^{ab}\frac{\partial \phi}{\partial x^b} + v^a\right) p\right] + g^{ab}\frac{\partial^2 p}{\partial v^a \partial v^b}$$

- Some schoolboy algebra shows that for $dp/dt = 0$

Second-Order Diffusions

- What is invariant measure of this diffusion?
- The Fokker-Planck equation for the diffusion follows as

$$\frac{dp}{dt} = - v^a \frac{\partial p}{\partial x^a} + \frac{\partial}{\partial v^a}\left[\left(\Gamma^a_{bc} v^b v^c + g^{ab}\frac{\partial \phi}{\partial x^b} + v^a\right) p\right] + g^{ab}\frac{\partial^2 p}{\partial v^a \partial v^b}$$

- Some schoolboy algebra shows that for $dp/dt = 0$

$$\frac{dp}{p} = \left[-\frac{1}{2}\frac{\partial g_{ab}}{\partial x^c} v^a v^b + \frac{\partial}{\partial x^c}\log\sqrt{|g|} - \frac{\partial \phi}{\partial x^x}\right] dx^c$$

Second-Order Diffusions

- What is invariant measure of this diffusion?
- The Fokker-Planck equation for the diffusion follows as

$$\frac{dp}{dt} = - \quad v^a \frac{\partial p}{\partial x^a} + \frac{\partial}{\partial v^a}\left[\left(\Gamma^a_{bc} v^b v^c + g^{ab}\frac{\partial \phi}{\partial x^b} + v^a\right) p\right] + g^{ab}\frac{\partial^2 p}{\partial v^a \partial v^b}$$

- Some schoolboy algebra shows that for $dp/dt = 0$

$$\frac{dp}{p} = \left[-\frac{1}{2}\frac{\partial g_{ab}}{\partial x^c} v^a v^b + \frac{\partial}{\partial x^c}\log\sqrt{|g|} - \frac{\partial \phi}{\partial x^x}\right] dx^c$$

- Therefore the above second-order SDE is satisfied by invariant densities $t \to \infty$ of the form

Second-Order Diffusions

- What is invariant measure of this diffusion?
- The Fokker-Planck equation for the diffusion follows as

$$\frac{dp}{dt} = - v^a \frac{\partial p}{\partial x^a} + \frac{\partial}{\partial v^a}\left[\left(\Gamma^a_{bc} v^b v^c + g^{ab}\frac{\partial \phi}{\partial x^b} + v^a\right) p\right] + g^{ab}\frac{\partial^2 p}{\partial v^a \partial v^b}$$

- Some schoolboy algebra shows that for $dp/dt = 0$

$$\frac{dp}{p} = \left[-\frac{1}{2}\frac{\partial g_{ab}}{\partial x^c} v^a v^b + \frac{\partial}{\partial x^c}\log\sqrt{|g|} - \frac{\partial \phi}{\partial x^x}\right] dx^c$$

- Therefore the above second-order SDE is satisfied by invariant densities $t \to \infty$ of the form

$$p(\mathbf{x}, \mathbf{v}) \propto \det(\mathbf{G}(\mathbf{x}))\exp\left(-\frac{1}{2}\mathbf{v}^{\mathsf{T}}\mathbf{G}(\mathbf{x})\mathbf{v} - \phi(\mathbf{x})\right)$$

Second-Order Diffusions

- What is invariant measure of this diffusion?
- The Fokker-Planck equation for the diffusion follows as

$$\frac{dp}{dt} = - v^a \frac{\partial p}{\partial x^a} + \frac{\partial}{\partial v^a}\left[\left(\Gamma^a_{bc} v^b v^c + g^{ab}\frac{\partial \phi}{\partial x^b} + v^a\right) p\right] + g^{ab}\frac{\partial^2 p}{\partial v^a \partial v^b}$$

- Some schoolboy algebra shows that for $dp/dt = 0$

$$\frac{dp}{p} = \left[-\frac{1}{2}\frac{\partial g_{ab}}{\partial x^c} v^a v^b + \frac{\partial}{\partial x^c}\log\sqrt{|g|} - \frac{\partial \phi}{\partial x^x}\right] dx^c$$

- Therefore the above second-order SDE is satisfied by invariant densities $t \to \infty$ of the form

$$p(\mathbf{x}, \mathbf{v}) \propto \det(\mathbf{G}(\mathbf{x})) \exp\left(-\frac{1}{2}\mathbf{v}^{\mathsf{T}}\mathbf{G}(\mathbf{x})\mathbf{v} - \phi(\mathbf{x})\right)$$

- So for $\phi(\mathbf{x}) = -\log \pi(\mathbf{x}) + \frac{1}{2}\log\det(\mathbf{G}(\mathbf{x}))$ then it follows that marginally

$$p(\mathbf{x}) = \exp(-\phi(\mathbf{x})) = \pi(\mathbf{x})$$

- Numerical integration forms basis of MCMC scheme..... however

Second-Order Diffusions

Second-Order Diffusions

- Transform the system of SDE's from the configuration to phase space such that transformed variable $\mathbf{p}_t = \mathbf{G}(\mathbf{x}_t)\mathbf{v}_t$

Second-Order Diffusions

- Transform the system of SDE's from the configuration to phase space such that transformed variable $\mathbf{p}_t = \mathbf{G}(\mathbf{x}_t)\mathbf{v}_t$

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{v}_t dt \\
d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

Second-Order Diffusions

- Transform the system of SDE's from the configuration to phase space such that transformed variable $\mathbf{p}_t = \mathbf{G}(\mathbf{x}_t)\mathbf{v}_t$

$$\begin{aligned} d\mathbf{x}_t &= \mathbf{v}_t dt \\ d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t \end{aligned}$$

- is transformed to

Second-Order Diffusions

- Transform the system of SDE's from the configuration to phase space such that transformed variable $\mathbf{p}_t = \mathbf{G}(\mathbf{x}_t)\mathbf{v}_t$

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{v}_t dt \\
d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- is transformed to

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{G}^{-1}(\mathbf{x}_t)\mathbf{p}_t dt \\
d\mathbf{p}_t &= -\boldsymbol{\nu}(\mathbf{x}_t, \mathbf{p}_t)dt - \nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{p}_t dt + \sqrt{2\mathbf{G}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

Second-Order Diffusions

- Transform the system of SDE's from the configuration to phase space such that transformed variable $\mathbf{p}_t = \mathbf{G}(\mathbf{x}_t)\mathbf{v}_t$

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{v}_t dt \\
d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- is transformed to

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{G}^{-1}(\mathbf{x}_t)\mathbf{p}_t dt \\
d\mathbf{p}_t &= -\boldsymbol{\nu}(\mathbf{x}_t, \mathbf{p}_t)dt - \nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{p}_t dt + \sqrt{2\mathbf{G}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- where vector function $\boldsymbol{\nu}$ has elements $-\partial_i g^{ab} p_a p_b$

Second-Order Diffusions

- Transform the system of SDE's from the configuration to phase space such that transformed variable $\mathbf{p}_t = \mathbf{G}(\mathbf{x}_t)\mathbf{v}_t$

$$\begin{aligned} d\mathbf{x}_t &= \mathbf{v}_t dt \\ d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t \end{aligned}$$

- is transformed to

$$\begin{aligned} d\mathbf{x}_t &= \mathbf{G}^{-1}(\mathbf{x}_t)\mathbf{p}_t dt \\ d\mathbf{p}_t &= -\boldsymbol{\nu}(\mathbf{x}_t, \mathbf{p}_t)dt - \nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{p}_t dt + \sqrt{2\mathbf{G}(\mathbf{x}_t)}d\mathbf{W}_t \end{aligned}$$

- where vector function $\boldsymbol{\nu}$ has elements $-\partial_i g^{ab} p_a p_b$
- Invariant measure with density

$$p(\mathbf{x}, \mathbf{p}) \propto \frac{1}{\det(\mathbf{G}(\mathbf{x}))} \exp\left(-\frac{1}{2}\mathbf{p}^\mathsf{T}\mathbf{G}^{-1}(\mathbf{x})\mathbf{p} - \phi(\mathbf{x})\right)$$

Second-Order Diffusions

- Transform the system of SDE's from the configuration to phase space such that transformed variable $\mathbf{p}_t = \mathbf{G}(\mathbf{x}_t)\mathbf{v}_t$

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{v}_t dt \\
d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- is transformed to

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{G}^{-1}(\mathbf{x}_t)\mathbf{p}_t dt \\
d\mathbf{p}_t &= -\boldsymbol{\nu}(\mathbf{x}_t, \mathbf{p}_t)dt - \nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{p}_t dt + \sqrt{2\mathbf{G}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- where vector function $\boldsymbol{\nu}$ has elements $-\partial_i g^{ab} p_a p_b$
- Invariant measure with density

$$
p(\mathbf{x}, \mathbf{p}) \propto \frac{1}{\det(\mathbf{G}(\mathbf{x}))} \exp\left(-\frac{1}{2}\mathbf{p}^{\mathsf{T}}\mathbf{G}^{-1}(\mathbf{x})\mathbf{p} - \phi(\mathbf{x})\right)
$$

- Stochastic Hamiltonian on Manifold - Lie-Trotter Splitting of Hamiltonian (deterministic and stochastic OU Process) and using symplectic integrator - samples drawn from invariant measure via RMHMC

Second-Order Diffusions

- Transform the system of SDE's from the configuration to phase space such that transformed variable $\mathbf{p}_t = \mathbf{G}(\mathbf{x}_t)\mathbf{v}_t$

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{v}_t dt \\
d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- is transformed to

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{G}^{-1}(\mathbf{x}_t)\mathbf{p}_t dt \\
d\mathbf{p}_t &= -\boldsymbol{\nu}(\mathbf{x}_t, \mathbf{p}_t)dt - \nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{p}_t dt + \sqrt{2\mathbf{G}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- where vector function $\boldsymbol{\nu}$ has elements $-\partial_i g^{ab} p_a p_b$
- Invariant measure with density

$$
p(\mathbf{x}, \mathbf{p}) \propto \frac{1}{\det(\mathbf{G}(\mathbf{x}))} \exp\left(-\frac{1}{2}\mathbf{p}^{\mathsf{T}}\mathbf{G}^{-1}(\mathbf{x})\mathbf{p} - \phi(\mathbf{x})\right)
$$

- Stochastic Hamiltonian on Manifold - Lie-Trotter Splitting of Hamiltonian (deterministic and stochastic OU Process) and using symplectic integrator - samples drawn from invariant measure via RMHMC

- Proposals follow local geodesics on manifold

Second-Order Diffusions

Second-Order Diffusions

- Stochastic Hamiltonian on Manifold driven by conditional Ornstein-Uhlenbeck process

Second-Order Diffusions

- Stochastic Hamiltonian on Manifold driven by conditional Ornstein-Uhlenbeck process

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{G}^{-1}(\mathbf{x}_t)\mathbf{p}_t dt \\
d\mathbf{p}_t &= -\boldsymbol{\nu}(\mathbf{x}_t, \mathbf{p}_t)dt - \nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{p}_t dt + \sqrt{2\mathbf{G}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

Second-Order Diffusions

- Stochastic Hamiltonian on Manifold driven by conditional Ornstein-Uhlenbeck process

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{G}^{-1}(\mathbf{x}_t)\mathbf{p}_t dt \\
d\mathbf{p}_t &= -\boldsymbol{\nu}(\mathbf{x}_t, \mathbf{p}_t)dt - \nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{p}_t dt + \sqrt{2\mathbf{G}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- Employ symplectic integrator e.g. RMHMC for drift term

Second-Order Diffusions

- Stochastic Hamiltonian on Manifold driven by conditional Ornstein-Uhlenbeck process

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{G}^{-1}(\mathbf{x}_t)\mathbf{p}_t dt \\
d\mathbf{p}_t &= -\boldsymbol{\nu}(\mathbf{x}_t, \mathbf{p}_t)dt - \nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{p}_t dt + \sqrt{2\mathbf{G}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- Employ symplectic integrator e.g. RMHMC for drift term
- Note that OU diffusion process corresponds to ad-hoc partial momentum updating

Second-Order Diffusions

- Stochastic Hamiltonian on Manifold driven by conditional Ornstein-Uhlenbeck process

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{G}^{-1}(\mathbf{x}_t)\mathbf{p}_t dt \\
d\mathbf{p}_t &= -\boldsymbol{\nu}(\mathbf{x}_t, \mathbf{p}_t)dt - \nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{p}_t dt + \sqrt{2\mathbf{G}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- Employ symplectic integrator e.g. RMHMC for drift term
- Note that OU diffusion process corresponds to ad-hoc partial momentum updating
- Solid theoretical basis amenable to analysis of algorithm performance

Second-Order Diffusions

- Stochastic Hamiltonian on Manifold driven by conditional Ornstein-Uhlenbeck process

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{G}^{-1}(\mathbf{x}_t)\mathbf{p}_t dt \\
d\mathbf{p}_t &= -\boldsymbol{\nu}(\mathbf{x}_t, \mathbf{p}_t)dt - \nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{p}_t dt + \sqrt{2\mathbf{G}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- Employ symplectic integrator e.g. RMHMC for drift term
- Note that OU diffusion process corresponds to ad-hoc partial momentum updating
- Solid theoretical basis amenable to analysis of algorithm performance
- Phase space defined on co-tangent bundle with symplectic forms in place

Second-Order Diffusions

- Stochastic Hamiltonian on Manifold driven by conditional Ornstein-Uhlenbeck process

$$\begin{aligned} d\mathbf{x}_t &= \mathbf{G}^{-1}(\mathbf{x}_t)\mathbf{p}_t dt \\ d\mathbf{p}_t &= -\boldsymbol{\nu}(\mathbf{x}_t, \mathbf{p}_t)dt - \nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{p}_t dt + \sqrt{2\mathbf{G}(\mathbf{x}_t)}d\mathbf{W}_t \end{aligned}$$

- Employ symplectic integrator e.g. RMHMC for drift term
- Note that OU diffusion process corresponds to ad-hoc partial momentum updating
- Solid theoretical basis amenable to analysis of algorithm performance
- Phase space defined on co-tangent bundle with symplectic forms in place
- What of the configuration space?

Second-Order Diffusions

Second-Order Diffusions

- This SDE defined on co-tangent bundle - non-symplectic forms

Second-Order Diffusions

- This SDE defined on co-tangent bundle - non-symplectic forms

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{v}_t dt \\
d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

Second-Order Diffusions

- This SDE defined on co-tangent bundle - non-symplectic forms

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{v}_t dt \\
d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- Drift function can be seen to describe Lagrangian dynamics

Second-Order Diffusions

- This SDE defined on co-tangent bundle - non-symplectic forms

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{v}_t dt \\
d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- Drift function can be seen to describe Lagrangian dynamics
- OU diffusion driving overall process

Second-Order Diffusions

- This SDE defined on co-tangent bundle - non-symplectic forms

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{v}_t dt \\
d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- Drift function can be seen to describe Lagrangian dynamics
- OU diffusion driving overall process
- Construct explicit integrator for dynamics - with volume correction

Second-Order Diffusions

- This SDE defined on co-tangent bundle - non-symplectic forms

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{v}_t dt \\
d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- Drift function can be seen to describe Lagrangian dynamics
- OU diffusion driving overall process
- Construct explicit integrator for dynamics - with volume correction
- Preserves original Hamiltonian dynamics - high proposal acceptance rates

Second-Order Diffusions

- This SDE defined on co-tangent bundle - non-symplectic forms

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{v}_t dt \\
d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- Drift function can be seen to describe Lagrangian dynamics
- OU diffusion driving overall process
- Construct explicit integrator for dynamics - with volume correction
- Preserves original Hamiltonian dynamics - high proposal acceptance rates
- Speed improvement over implicit symplectic integrator between 2x to 10x

Second-Order Diffusions

- This SDE defined on co-tangent bundle - non-symplectic forms

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{v}_t dt \\
d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- Drift function can be seen to describe Lagrangian dynamics
- OU diffusion driving overall process
- Construct explicit integrator for dynamics - with volume correction
- Preserves original Hamiltonian dynamics - high proposal acceptance rates
- Speed improvement over implicit symplectic integrator between 2x to 10x Though non-symplectic geometry limits sampling - see The Geometry Foundations of Hamiltonian Monte Carlo `http://arxiv.org/pdf/1410.5110.pdf`

Second-Order Diffusions

- This SDE defined on co-tangent bundle - non-symplectic forms

$$
\begin{aligned}
d\mathbf{x}_t &= \mathbf{v}_t dt \\
d\mathbf{v}_t &= -\gamma(\mathbf{x}_t, \mathbf{v}_t)dt - \mathbf{G}^{-1}(\mathbf{x}_t)\nabla_{\mathbf{x}_t}\phi(\mathbf{x}_t)dt - \mathbf{v}_t dt + \sqrt{2\mathbf{G}^{-1}(\mathbf{x}_t)}d\mathbf{W}_t
\end{aligned}
$$

- Drift function can be seen to describe Lagrangian dynamics
- OU diffusion driving overall process
- Construct explicit integrator for dynamics - with volume correction
- Preserves original Hamiltonian dynamics - high proposal acceptance rates
- Speed improvement over implicit symplectic integrator between 2x to 10x
  Though non-symplectic geometry limits sampling - see The Geometry Foundations of Hamiltonian Monte Carlo
  `http://arxiv.org/pdf/1410.5110.pdf`
- Markov chain Monte Carlo from Lagrangian Dynamics
  Journal of Comp.Graph.Stats, 2014

Gaussian Mixture Model

- Univariate finite mixture model

$$p(x|\mu, \sigma^2) = 0.7 \times \mathcal{N}(x|0, \sigma^2) + 0.3 \times \mathcal{N}(x|\mu, \sigma^2)$$

Gaussian Mixture Model

- Univariate finite mixture model

$$p(x|\mu, \sigma^2) = 0.7 \times \mathcal{N}(x|0, \sigma^2) + 0.3 \times \mathcal{N}(x|\mu, \sigma^2)$$

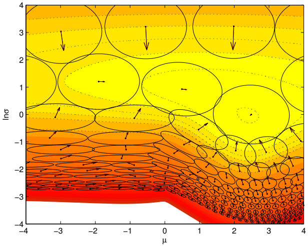
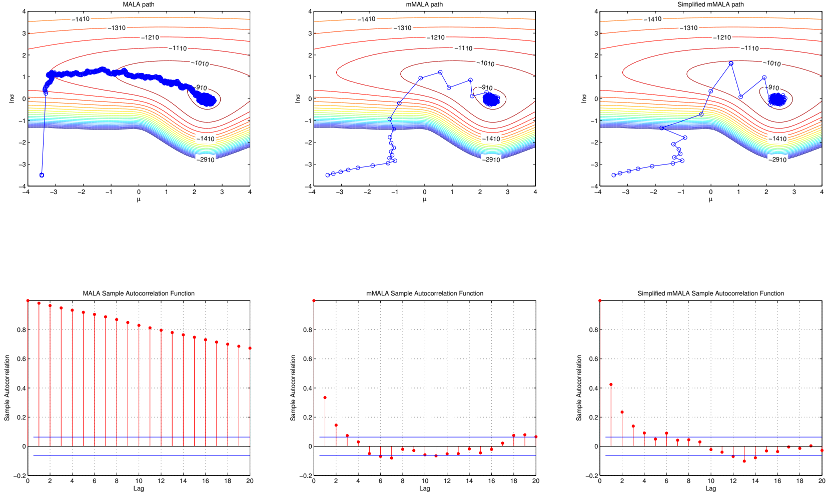
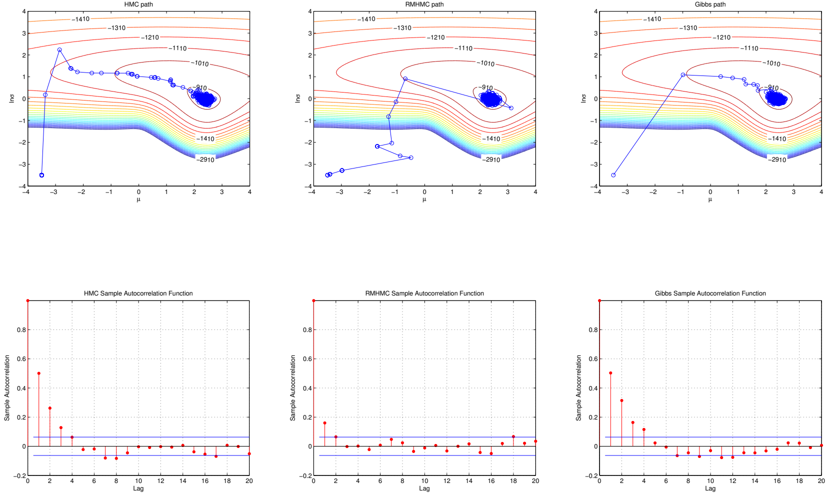


Figure : Arrows correspond to the gradients and ellipses to the inverse metric tensor. Dashed lines are isocontours of the joint log density

Gaussian Mixture Model



Gaussian Mixture Model



Log-Gaussian Cox Point Process with Latent Field

- The joint density for Poisson counts and latent Gaussian field

$$p(\mathbf{y}, \mathbf{x}|\mu, \sigma, \beta) \propto \prod_{i,j}^{64} \exp\{y_{i,j} x_{i,j} - m \exp(x_{i,j})\} \exp(-(\mathbf{x} - \mu \mathbf{1})^{\mathsf{T}} \Sigma_{\theta}^{-1} (\mathbf{x} - \mu \mathbf{1})/2)$$

Log-Gaussian Cox Point Process with Latent Field

- The joint density for Poisson counts and latent Gaussian field

$$p(\mathbf{y}, \mathbf{x} | \mu, \sigma, \beta) \propto \prod_{i,j}^{64} \exp\{y_{i,j} x_{i,j} - m \exp(x_{i,j})\} \exp(-(\mathbf{x} - \mu \mathbf{1})^{\mathsf{T}} \Sigma_{\theta}^{-1} (\mathbf{x} - \mu \mathbf{1})/2)$$

- Metric tensors

$$\begin{aligned} \mathbf{G}(\boldsymbol{\theta})_{i,j} &= \frac{1}{2} trace \left(\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\boldsymbol{\theta}}}{\partial \theta_i} \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\boldsymbol{\theta}}}{\partial \theta_j} \right) \\ \mathbf{G}(\mathbf{x}) &= \Lambda + \Sigma_{\boldsymbol{\theta}}^{-1} \end{aligned}$$

where Λ is diagonal with elements $m \exp(\mu + (\Sigma_{\boldsymbol{\theta}})_{i,i})$

Log-Gaussian Cox Point Process with Latent Field

- The joint density for Poisson counts and latent Gaussian field

$$p(\mathbf{y},\mathbf{x}|\mu,\sigma,\beta) \propto \prod_{i,j}^{64} \exp\{y_{i,j}x_{i,j} - m\exp(x_{i,j})\}\exp(-(\mathbf{x}-\mu\mathbf{1})^{\mathsf{T}}\Sigma_{\theta}^{-1}(\mathbf{x}-\mu\mathbf{1})/2)$$

- Metric tensors

$$\begin{aligned}\mathbf{G}(\boldsymbol{\theta})_{i,j} &= \frac{1}{2}\,\mathit{trace}\left(\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}\frac{\partial\boldsymbol{\Sigma}_{\boldsymbol{\theta}}}{\partial\theta_i}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}\frac{\partial\boldsymbol{\Sigma}_{\boldsymbol{\theta}}}{\partial\theta_j}\right)\\ \mathbf{G}(\mathbf{x}) &= \Lambda + \Sigma_{\boldsymbol{\theta}}^{-1}\end{aligned}$$

  where Λ is diagonal with elements $m\exp(\mu + (\Sigma_{\boldsymbol{\theta}})_{i,i})$

- Latent field metric tensor defining flat manifold is 4096×4096, $\mathcal{O}(N^3)$ obtained from parameter conditional

Log-Gaussian Cox Point Process with Latent Field

- The joint density for Poisson counts and latent Gaussian field

$$p(\mathbf{y}, \mathbf{x}|\mu, \sigma, \beta) \propto \prod_{i,j}^{64} \exp\{y_{i,j}x_{i,j} - m\exp(x_{i,j})\} \exp(-(\mathbf{x}-\mu\mathbf{1})^{\mathsf{T}}\Sigma_{\theta}^{-1}(\mathbf{x}-\mu\mathbf{1})/2)$$

- Metric tensors

$$\begin{aligned}
\mathbf{G}(\boldsymbol{\theta})_{i,j} &= \frac{1}{2} trace\left(\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}\frac{\partial\boldsymbol{\Sigma}_{\boldsymbol{\theta}}}{\partial\theta_i}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}\frac{\partial\boldsymbol{\Sigma}_{\boldsymbol{\theta}}}{\partial\theta_j}\right) \\
\mathbf{G}(\mathbf{x}) &= \Lambda + \Sigma_{\boldsymbol{\theta}}^{-1}
\end{aligned}$$

  where Λ is diagonal with elements $m\exp(\mu + (\Sigma_{\boldsymbol{\theta}})_{i,i})$

- Latent field metric tensor defining flat manifold is 4096×4096, $\mathcal{O}(N^3)$ obtained from parameter conditional

- MALA requires transformation of latent field to sample - with separate tuning in transient and stationary phases of Markov chain

RMHMC for Log-Gaussian Cox Point Processes

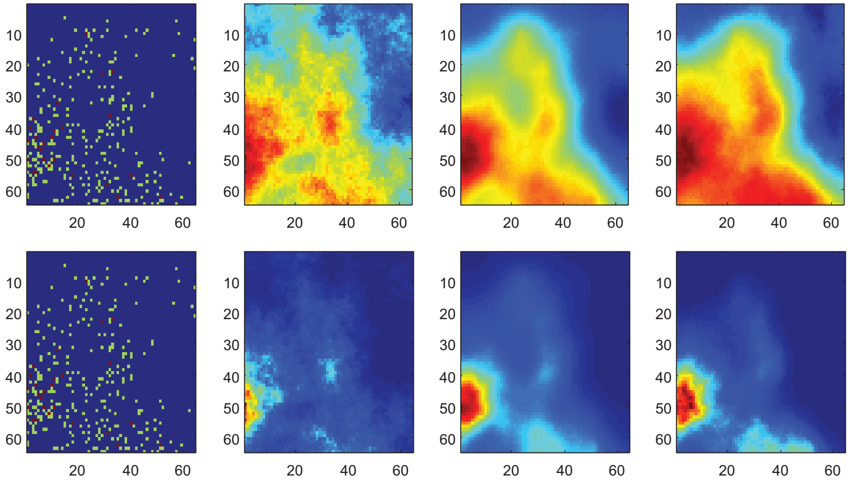


Figure : Data, Latent Field, Poisson Mean Field

RMHMC for Log-Gaussian Cox Point Processes

Table : Sampling the latent variables of a Log-Gaussian Cox Process - Comparison of sampling methods

Method	Time	ESS (Min, Med, Max)	s/Min ESS	Rel. Speed
MALA (Transient)	31,577	(3, 8, 50)	10,605	$\times 1$
MALA (Stationary)	31,118	(4, 16, 80)	7836	$\times 1.35$
mMALA	634	(26, 84, 174)	24.1	$\times 440$
RMHMC	2936	(1951, 4545, 5000)	1.5	$\times 7070$

Illustrative Heat Conduction Problem

- Heat conduction problem governed by an elliptic partial differential equation in the open and bounded domain Ω:

Illustrative Heat Conduction Problem

- Heat conduction problem governed by an elliptic partial differential equation in the open and bounded domain Ω:

$$
\begin{aligned}
-\nabla \cdot (e^{u} \nabla w) &= 0 \quad in \quad \Omega \\
-e^{u} \nabla w \cdot \mathrm{n} &= u \;\, \mathrm{Bi} \quad on \quad \partial\Omega / \, \Gamma_R \\
-e^{u} \nabla w \cdot \mathrm{n} &= -1 \quad on \quad \Gamma_R
\end{aligned}
$$

Illustrative Heat Conduction Problem

- Heat conduction problem governed by an elliptic partial differential equation in the open and bounded domain Ω:

$$
\begin{aligned}
-\nabla \cdot (e^{u} \nabla w) &= 0 \quad \mathit{in} \quad \Omega \\
-e^{u} \nabla w \cdot \mathrm{n} &= u \ \mathrm{Bi} \quad \mathit{on} \quad \partial\Omega / \, \Gamma_R \\
-e^{u} \nabla w \cdot \mathrm{n} &= -1 \quad \mathit{on} \quad \Gamma_R
\end{aligned}
$$

- Forward state is w and u the logarithm of distributed thermal conductivity on Ω, n the unit outward normal on Ω, and Bi the Biot number.

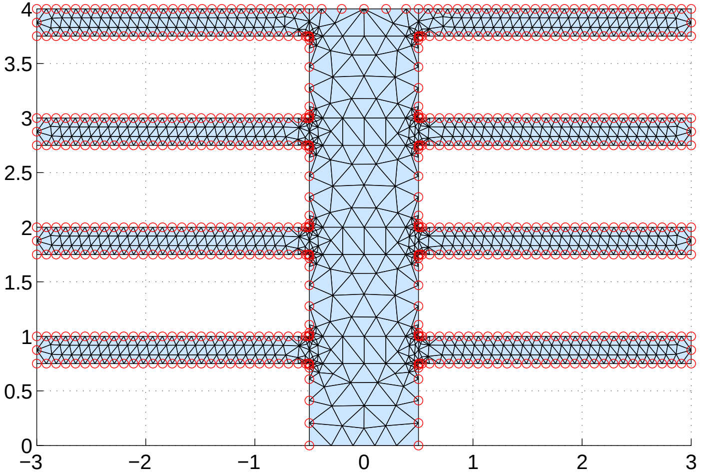
Illustrative Heat Conduction Problem

- Heat conduction problem governed by an elliptic partial differential equation in the open and bounded domain Ω:

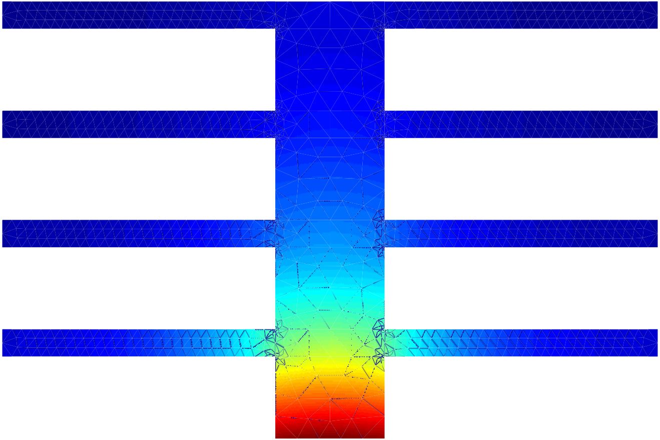
$$
\begin{aligned}
-\nabla \cdot (e^{u} \nabla w) &= 0 \quad in \quad \Omega \\
-e^{u} \nabla w \cdot \mathrm{n} &= u \ \mathrm{Bi} \quad on \quad \partial\Omega / \, \Gamma_R \\
-e^{u} \nabla w \cdot \mathrm{n} &= -1 \quad on \quad \Gamma_R
\end{aligned}
$$

- Forward state is w and u the logarithm of distributed thermal conductivity on Ω, n the unit outward normal on Ω, and Bi the Biot number.

Illustrative Heat Conduction Problem



Illustrative Heat Conduction Problem

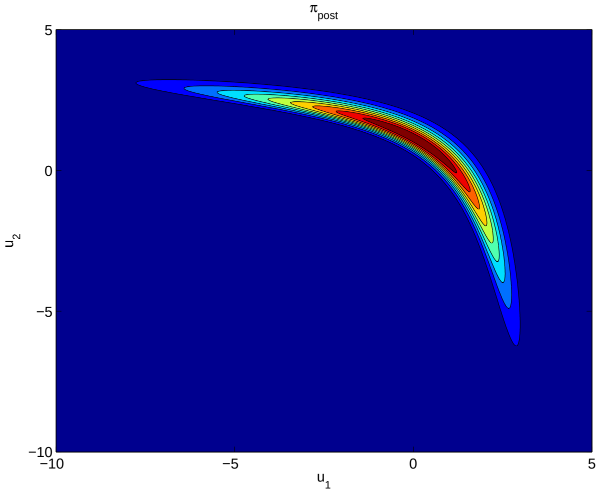


Illustrative Heat Conduction Problem

- Take one finite element discretisation of domain and one observation at leftmost boundary

Illustrative Heat Conduction Problem

- Take one finite element discretisation of domain and one observation at leftmost boundary
- Consider induced bivariate posterior for varying forms of prior Gaussian measure



Hamiltonian on Riemann Manifold

The dynamics for k-th component of $\mathbf{u}$ is given by Hamiltons equations

$$\frac{d\mathbf{u}_k}{dt} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}_k} = \left(\mathbf{G}(\mathbf{u})^{-1}\mathbf{p}\right)_k$$

$$\frac{d\mathbf{p}_k}{dt} = -\frac{\partial \mathcal{H}}{\partial \mathbf{u}_k} = -\nabla_k J(\mathbf{u}) - \frac{1}{2} Tr\left[\mathbf{G}(\mathbf{u})^{-1}\frac{\partial\, \mathbf{G}(\mathbf{u})}{\partial \mathbf{u}_k}\right] + \frac{1}{2}\mathbf{p}^T \mathbf{G}(\mathbf{u})^{-1}\frac{\partial\, \mathbf{G}(\mathbf{u})}{\partial \mathbf{u}_k}\mathbf{G}(\mathbf{u})^{-1}\mathbf{p}$$

Gradient

Gradient

$$\langle \nabla \mathcal{J}(u), \tilde{u} \rangle = \int_\Omega \tilde{u} e^u \nabla w \cdot \nabla \lambda \, d\Omega,$$

First Order Forward

$$\int_\Omega e^u \nabla w \cdot \nabla \hat{\lambda} \, d\Omega + \int_{\partial\Omega \setminus \Gamma_R} Bi \, w \hat{\lambda} \, ds = \int_{\Gamma_R} \hat{\lambda} \, ds,$$

First Order Adjoint

$$\int_\Omega e^u \nabla \lambda \cdot \nabla \hat{w} \, d\Omega + \int_{\partial\Omega \setminus \Gamma_R} Bi \, \lambda \hat{w} \, ds = -\frac{1}{\sigma^2} \sum_{j=1}^{K} \left(w(\mathbf{x}_j) - d_j \right) \hat{w}(\mathbf{x}_j),$$

Hamiltonian on Riemann Manifold

The dynamics for k-th component of u is given by Hamiltons equations

$$\frac{du_k}{dt} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}_k} = \left(\mathbf{G}(u)^{-1}\mathbf{p}\right)_k$$

$$\frac{d\mathbf{p}_k}{dt} = -\frac{\partial \mathcal{H}}{\partial u_k} = -\nabla_k J(u) - \frac{1}{2} Tr\left[\mathbf{G}(u)^{-1}\frac{\partial\, \mathbf{G}(u)}{\partial u_k}\right] + \frac{1}{2}\mathbf{p}^T \mathbf{G}(u)^{-1}\frac{\partial\, \mathbf{G}(u)}{\partial u_k}\mathbf{G}(u)^{-1}\mathbf{p}$$

Hessian/Fisher Information Matrix

Fisher Matrix-Vector Product

$$\left\langle \langle G(u), \tilde{u} \rangle , u^2 \right\rangle = \int_\Omega \tilde{u} e^u \nabla w \cdot \nabla \tilde{\lambda}^2 \, d\Omega,$$

Second Order Forward

$$\int_\Omega e^u \nabla w^2 \cdot \nabla \hat{\lambda} \, d\Omega + \int_{\partial\Omega \backslash \Gamma_R} Bi \, w^2 \hat{\lambda} \, ds = - \int_\Omega u^2 e^u \nabla w \cdot \nabla \hat{\lambda} \, d\Omega,$$

Second Order Adjoint

$$\int_\Omega e^u \nabla \tilde{\lambda}^2 \cdot \nabla \hat{w} \, d\Omega + \int_{\partial\Omega \backslash \Gamma_R} Bi \, \tilde{\lambda}^2 \hat{w} \, ds = -\frac{1}{\sigma^2} \sum_{j=1}^{K} w^2 (\mathbf{x}_j) \, \hat{w} (\mathbf{x}_j) \, .$$

Hamiltonian on Riemann Manifold

The dynamics for k-th component of u is given by Hamiltons equations

$$\frac{du_k}{dt} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}_k} = \left(\mathbf{G}(u)^{-1}\mathbf{p}\right)_k$$

$$\frac{d\mathbf{p}_k}{dt} = -\frac{\partial \mathcal{H}}{\partial u_k} = -\nabla_k J(u) - \frac{1}{2} Tr\left[\mathbf{G}(u)^{-1}\frac{\partial \mathbf{G}(u)}{\partial u_k}\right] + \frac{1}{2}\mathbf{p}^T \mathbf{G}(u)^{-1}\frac{\partial \mathbf{G}(u)}{\partial u_k}\mathbf{G}(u)^{-1}\mathbf{p}$$

Derivative of Fisher Information Matrix

Derivative of Fisher Matrix-Matrix product

$$\left\langle\left\langle\left\langle T\left(u\right),\tilde{u}\right\rangle,u^2\right\rangle,u^3\right\rangle := \left\langle\nabla\left\langle\left\langle G\left(u\right),\tilde{u}\right\rangle,u^2\right\rangle,u^3\right\rangle$$
$$= \int_\Omega \tilde{u}u^3 e^u \nabla w \cdot \nabla\tilde{\lambda^2}\,d\Omega + \int_\Omega \tilde{u}e^u \nabla w^3 \cdot \nabla\tilde{\lambda^2}\,d\Omega + \int_\Omega \tilde{u}e^u \nabla w \cdot \nabla\lambda^{2,3}\,d\Omega$$

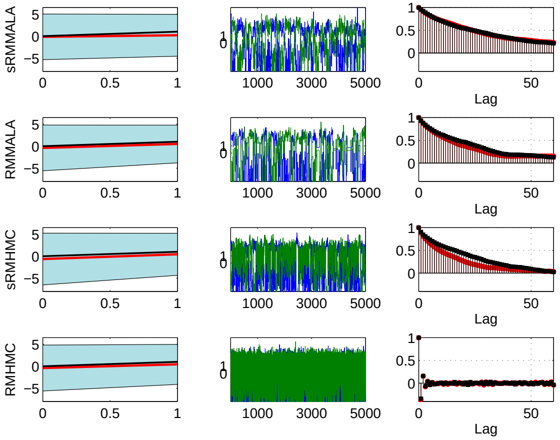
Third Order Forward

$$\int_\Omega e^u \nabla w^3 \cdot \nabla\hat{\lambda}\,d\Omega + \int_{\partial\Omega\setminus\Gamma_R} Bi\,w^3\hat{\lambda}\,ds = -\int_\Omega u^3 e^u \nabla w \cdot \nabla\hat{\lambda}\,d\Omega.$$

Third Order Adjoint

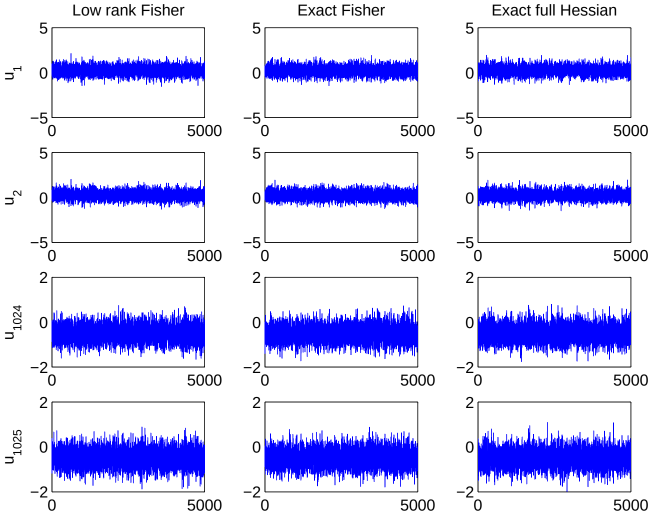
$$\int_\Omega e^u \nabla\lambda^{2,3} \cdot \nabla\hat{w}\,d\Omega + \int_{\partial\Omega\setminus\Gamma_R} Bi\,\lambda^{2,3}\hat{w}\,ds = -\frac{1}{\sigma^2}\sum_{j=1}^{K} w^{2,3}\left(\mathbf{x}_j\right)\hat{w}\left(\mathbf{x}_j\right)$$
$$-\int_\Omega u^3 e^u \nabla\tilde{\lambda^2} \cdot \nabla\hat{w}\,d\Omega,$$

Two-parameter Case

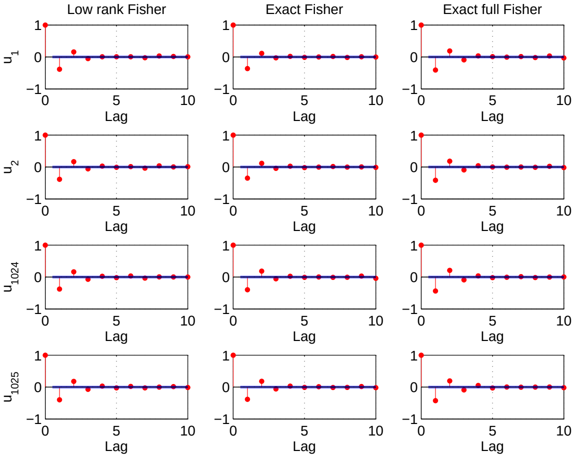


- $\varepsilon = 0.7$ for simRMMALA and RMMALA
- $\varepsilon = 0.02$, $L = 100$ for simRMHMC and RMHMC

1025-parameter Case



1025-parameter Case



Bui-Thanh, T., and Girolami, M., *Solving Large-scale PDE using Riemann Manifold Hamiltonian MCMC*, **Inverse Problems, 30, 92014) 114014, IoP Publishing**.

Conclusion and Discussion

- Geometry of statistical models harnessed in Monte Carlo methods

Conclusion and Discussion

- Geometry of statistical models harnessed in Monte Carlo methods
  - Diffusions that respect structure and curvature of space - Manifold MALA

Conclusion and Discussion

- Geometry of statistical models harnessed in Monte Carlo methods
  - Diffusions that respect structure and curvature of space - Manifold MALA
  - Geodesic flows on model manifold - RMHMC - generalisation of HMC

Conclusion and Discussion

- Geometry of statistical models harnessed in Monte Carlo methods
  - Diffusions that respect structure and curvature of space - Manifold MALA
  - Geodesic flows on model manifold - RMHMC - generalisation of HMC
  - Assessed on correlated & high-dimensional latent variable models

Conclusion and Discussion

- Geometry of statistical models harnessed in Monte Carlo methods
  - Diffusions that respect structure and curvature of space - Manifold MALA
  - Geodesic flows on model manifold - RMHMC - generalisation of HMC
  - Assessed on correlated & high-dimensional latent variable models
  - Promising capability of methodology

Conclusion and Discussion

- Geometry of statistical models harnessed in Monte Carlo methods
  - Diffusions that respect structure and curvature of space - Manifold MALA
  - Geodesic flows on model manifold - RMHMC - generalisation of HMC
  - Assessed on correlated & high-dimensional latent variable models
  - Promising capability of methodology
- Ongoing development

Conclusion and Discussion

- Geometry of statistical models harnessed in Monte Carlo methods
  - Diffusions that respect structure and curvature of space - Manifold MALA
  - Geodesic flows on model manifold - RMHMC - generalisation of HMC
  - Assessed on correlated & high-dimensional latent variable models
  - Promising capability of methodology
- Ongoing development
  - Theoretical analysis of convergence

Conclusion and Discussion

- Geometry of statistical models harnessed in Monte Carlo methods
  - Diffusions that respect structure and curvature of space - Manifold MALA
  - Geodesic flows on model manifold - RMHMC - generalisation of HMC
  - Assessed on correlated & high-dimensional latent variable models
  - Promising capability of methodology
- Ongoing development
  - Theoretical analysis of convergence
  - Geodesic Monte Carlo on Embedded Manifolds
    `http://arxiv.org/pdf/1301.6064.pdf`

Conclusion and Discussion

- Geometry of statistical models harnessed in Monte Carlo methods
  - Diffusions that respect structure and curvature of space - Manifold MALA
  - Geodesic flows on model manifold - RMHMC - generalisation of HMC
  - Assessed on correlated & high-dimensional latent variable models
  - Promising capability of methodology
- Ongoing development
  - Theoretical analysis of convergence
  - Geodesic Monte Carlo on Embedded Manifolds
    `http://arxiv.org/pdf/1301.6064.pdf`
  - Investigate alternative manifold structures

Conclusion and Discussion

- Geometry of statistical models harnessed in Monte Carlo methods
  - Diffusions that respect structure and curvature of space - Manifold MALA
  - Geodesic flows on model manifold - RMHMC - generalisation of HMC
  - Assessed on correlated & high-dimensional latent variable models
  - Promising capability of methodology
- Ongoing development
  - Theoretical analysis of convergence
  - Geodesic Monte Carlo on Embedded Manifolds
    `http://arxiv.org/pdf/1301.6064.pdf`
  - Investigate alternative manifold structures
  - Design and effect of metric and connection

Conclusion and Discussion

- Geometry of statistical models harnessed in Monte Carlo methods
  - Diffusions that respect structure and curvature of space - Manifold MALA
  - Geodesic flows on model manifold - RMHMC - generalisation of HMC
  - Assessed on correlated & high-dimensional latent variable models
  - Promising capability of methodology
- Ongoing development
  - Theoretical analysis of convergence
  - Geodesic Monte Carlo on Embedded Manifolds
    `http://arxiv.org/pdf/1301.6064.pdf`
  - Investigate alternative manifold structures
  - Design and effect of metric and connection
  - Optimality of Hamiltonian flows as local geodesics

Conclusion and Discussion

- Geometry of statistical models harnessed in Monte Carlo methods
  - Diffusions that respect structure and curvature of space - Manifold MALA
  - Geodesic flows on model manifold - RMHMC - generalisation of HMC
  - Assessed on correlated & high-dimensional latent variable models
  - Promising capability of methodology
- Ongoing development
  - Theoretical analysis of convergence
  - Geodesic Monte Carlo on Embedded Manifolds `http://arxiv.org/pdf/1301.6064.pdf`
  - Investigate alternative manifold structures
  - Design and effect of metric and connection
  - Optimality of Hamiltonian flows as local geodesics
  - Surrogate geometries based on emulators for PDEs

Conclusion and Discussion

- Geometry of statistical models harnessed in Monte Carlo methods
  - Diffusions that respect structure and curvature of space - Manifold MALA
  - Geodesic flows on model manifold - RMHMC - generalisation of HMC
  - Assessed on correlated & high-dimensional latent variable models
  - Promising capability of methodology
- Ongoing development
  - Theoretical analysis of convergence
  - Geodesic Monte Carlo on Embedded Manifolds
    `http://arxiv.org/pdf/1301.6064.pdf`
  - Investigate alternative manifold structures
  - Design and effect of metric and connection
  - Optimality of Hamiltonian flows as local geodesics
  - Surrogate geometries based on emulators for PDEs
- No silver bullet or cure all - new powerful methodology for MC toolkit

Funding Acknowledgments

- EPSRC Established Research Career Fellowship
- EPSRC Programme Grant - Enabling Quantification of Uncertainty in Inverse Problems
- Royal Society Wolfson Research Merit Award