Geodesic slice sampling on Riemannian manifolds

Mareike Hasenpflug

University of Passau

April 2024

joint work with Alain Durmus, Samuel Gruffaz, Michael Habeck, Shantanu Kodgirwar and Daniel Rudolf



Manifolds

Definition

A *d*-dimensional *manifold* M is a second countable Hausdorff space such that for all $x \in M$ there exists a homeomorphism

$$\varphi: \mathsf{M} \supseteq \mathsf{U} \to \mathsf{U}' \subseteq \mathbb{R}^d$$
 and $x \in \mathsf{U}$ open.

The pair (U, φ) is called a *coordinate neighbourhood* of M.



Manifolds

Definition

A *d*-dimensional *manifold* M is a second countable Hausdorff space such that for all $x \in M$ there exists a homeomorphism

$$\varphi: \mathsf{M} \supseteq \mathsf{U} \to \mathsf{U}' \subseteq \mathbb{R}^d$$
 and $x \in \mathsf{U}$ open.

The pair (U, φ) is called a *coordinate neighbourhood* of M.



Manifolds

Definition

A *d*-dimensional *manifold* M is a second countable Hausdorff space such that for all $x \in M$ there exists a homeomorphism

$$arphi: \mathsf{M}\supseteq\mathsf{U}
ightarrow\mathsf{U}'\subseteq\mathbb{R}^d$$
 and $x\in\mathsf{U}$ open.

The pair (U, φ) is called a *coordinate neighbourhood* of M.

Aim: Generate samples from a distribution on a manifold.

Distributions on manifolds

We can incorporate constraints or dependencies directly into a model via manifold state spaces.

Example (Mantoux et al. 2021)

Model adjacency matrix $A \in \mathbb{R}^{d \times d}$ of an undirected graph through its eigendecomposition:

$$A = X \operatorname{diag}(a) X^{\top} + \operatorname{noise}, \qquad a \in \mathbb{R}^k, X \in \operatorname{Stiefel}(d, k).$$

 Applying e.g. Bayesian inference or MCMC-SAEM requires sampling from distributions on manifolds.

► To every point x ∈ M of a smooth manifold we can attach the tangent space T_xM, which is an ℝ-vector space.



- ► To every point x ∈ M of a smooth manifold we can attach the tangent space T_xM, which is an ℝ-vector space.
- ► The Riemannian metric g of a Riemannian manifold M induces an inner product g_x on T_xM for all x ∈ M.



- ► To every point x ∈ M of a smooth manifold we can attach the tangent space T_xM, which is an ℝ-vector space.
- ► The Riemannian metric g of a Riemannian manifold M induces an inner product g_x on T_xM for all x ∈ M.



- ► To every point x ∈ M of a smooth manifold we can attach the tangent space T_xM, which is an ℝ-vector space.
- ► The Riemannian metric g of a Riemannian manifold M induces an inner product g_x on T_xM for all x ∈ M.
- A Riemannian manifold M can be equipped with a generalisation of the Lebesgue measure, called the Riemannian measure v_g.



Problem description

Let M be a Riemannian manifold with Borel- σ -algebra $\mathcal{B}(M)$ and Riemannian measure $\nu_{\mathfrak{g}}$.

Consider a lower semi-continuous function

$$p:\mathsf{M} o (0,\infty) \qquad ext{with} \qquad Z:=\int_\mathsf{M} p(x)
u_\mathfrak{g}(\mathrm{d} x)\in (0,\infty).$$

Goal: Sample from

$$\pi(\mathrm{d} x) := \frac{1}{Z} p(x) \nu_{\mathfrak{g}}(\mathrm{d} x).$$

Approach: Simulate a Markov chain that has invariant distribution π .

Define the superlevel sets

$$L(t) := \{x \in M \mid p(x) > t\}, \quad t \ge 0.$$

The following algorithm is reversible w.r.t. π :

Input: $x \in M$ 1: Draw a level t from Unif(0, p(x)). 2: Draw y from $\nu_{\mathfrak{g}}(L(t))^{-1}\nu_{\mathfrak{g}}|_{L(t)}$. Output: $y \in M$



Define the superlevel sets

$$L(t) := \{x \in M \mid p(x) > t\}, \quad t \ge 0.$$

The following algorithm is reversible w.r.t. π :

Input: $x \in M$

1: Draw a level t from Unif(0, p(x)). 2: Draw y from $\nu_{\mathfrak{g}}(L(t))^{-1}\nu_{\mathfrak{g}}|_{L(t)}$. **Output:** $y \in M$



Define the superlevel sets

$$L(t) := \{x \in M \mid p(x) > t\}, \quad t \ge 0.$$

The following algorithm is reversible w.r.t. π :

Input: $x \in M$ 1: Draw a level t from Unif(0, p(x)). 2: Draw y from $\nu_{\mathfrak{g}}(L(t))^{-1}\nu_{\mathfrak{g}}|_{L(t)}$. Output: $y \in M$



Define the superlevel sets

$$L(t) := \{x \in M \mid p(x) > t\}, \quad t \ge 0.$$

The following algorithm is reversible w.r.t. π :

```
Input: x \in M

1: Draw a level t from \text{Unif}(0, p(x)).

2: Draw y from \nu_{\mathfrak{g}}(L(t))^{-1}\nu_{\mathfrak{g}}|_{L(t)}.

Output: y \in M
```



Define the superlevel sets

$$L(t) := \{x \in M \mid p(x) > t\}, \quad t \ge 0.$$

The following algorithm is reversible w.r.t. π :

Input: $x \in M$ 1: Draw a level t from Unif(0, p(x)). 2: Draw y from $\nu_{\mathfrak{g}}(L(t))^{-1}\nu_{\mathfrak{g}}|_{L(t)}$. Output: $y \in M$



Define the superlevel sets

$$L(t):=\{x\in \mathsf{M}\mid p(x)>t\},\qquad t\geq 0.$$

The following algorithm is reversible w.r.t. π :

Input: $x \in M$ 1: Draw a level t from Unif(0, p(x)). 2: Draw y from $\nu_{g}(L(t))^{-1}\nu_{g}|_{L(t)}$. Output: $y \in M$



Define the superlevel sets

$$L(t):=\{x\in \mathsf{M}\mid p(x)>t\},\qquad t\geq 0.$$

The following algorithm is reversible w.r.t. π :

Input: $x \in M$ 1: Draw a level t from Unif(0, p(x)). 2: Draw y from $\nu_{\mathfrak{g}}(L(t))^{-1}\nu_{\mathfrak{g}}|_{L(t)}$. Output: $y \in M$



Define the superlevel sets

$$L(t) := \{x \in M \mid p(x) > t\}, \quad t \ge 0.$$

The following algorithm is reversible w.r.t. π :

```
Input: x \in M

1: Draw a level t from \text{Unif}(0, p(x)).

2: Draw y from \nu_{\mathfrak{g}}(L(t))^{-1}\nu_{\mathfrak{g}}|_{L(t)}.

Output: y \in M
```

• "Spectral gap of the slice sampler only depends on level set function $t \mapsto \nu_{\mathfrak{g}}(L(t))$."¹

¹Natarovskii, Rudolf, and Sprungk 2021; Schär 2023; Rudolf and Schär 2024.

Mareike Hasenpflug (Uni. Passau) Geodesic slice sampling on Riemannian manifolds

Hybrid slice sampling

Problem: How to sample from $\nu_{\mathfrak{g}}(L(t))^{-1}\nu_{\mathfrak{g}}|_{L(t)}$?

Input: $x \in M$ 1: Draw a level t from Unif(0, p(x)). 2: Draw y from $\nu_{\mathfrak{g}}(L(t))^{-1}\nu_{\mathfrak{g}}|_{L(t)}$. Output: $y \in M$

Hybrid slice sampling

Problem: How to sample from $\nu_{\mathfrak{g}}(L(t))^{-1}\nu_{\mathfrak{g}}|_{L(t)}$?

Input: $x \in M$ 1: Draw a level t from Unif(0, p(x)). 2: Draw y from Markov kernel invariant w.r.t. $\nu_{\mathfrak{g}}(L(t))^{-1}\nu_{\mathfrak{g}}|_{L(t)}$. Output: $y \in M$

Approach: Replace step 2 by Markov kernel with inv. distribution $\nu_{\mathfrak{g}}(L(t))^{-1}\nu_{\mathfrak{g}}|_{L(t)}$.

Special case: The sphere

Define (d-2)-dimensional great subsphere with pole $x\in \mathbb{S}^{d-1}$ as

$$\mathbb{S}_{x}^{d-2} := \{ y \in \mathbb{S}^{d-1} \mid y^{T} x = 0 \}.$$

For each $x \in \mathbb{S}^{d-1}$ and $v \in \mathbb{S}^{d-2}_x$ there exists a unique great circle with $\gamma_{(x,v)}(0) = x$ and velocity vector v:



▶ Apply techniques² from 1-dim. slice sampling to intersection of level set and great circle.



²Neal 2003.

1: Draw t from Unif(0, p(x)). 2: Draw v from Unif (\mathbb{S}^{d-2}) . 3: Draw θ from Unif $(0, 2\pi)$. 4: Set $\theta_{\min} = \theta - 2\pi$ and $\theta_{\max} = \theta$. 5: repeat: Draw θ from Unif $(\theta_{\min}, \theta_{\max})$. 6: Set $y = \cos(\theta)x + \sin(\theta)y$. 7: $if\theta < 0$ 8: 9: $\theta_{\min} = \theta$ else 10: $\theta_{\max} = \theta$ 11: 12: **until** p(y) > t



- 1: Draw t from Unif(0, p(x)).
- 2: Draw v from $\text{Unif}(\mathbb{S}^{d-2}_x)$.
- 3: Draw θ from Unif $(0, 2\pi)$.
- 4: Set $\theta_{\min} = \theta 2\pi$ and $\theta_{\max} = \theta$.
- 5: repeat:
- 6: Draw θ from Unif $(\theta_{\min}, \theta_{\max})$.
- 7: Set $y = \cos(\theta)x + \sin(\theta)v$.
- 8: $\mathbf{if}\theta < \mathbf{0}$
- 9: $\theta_{\min} = \theta$
- 10: else
- 11: $\theta_{\max} = \theta$
- 12: **until** p(y) > t



- 1: Draw t from Unif(0, p(x)).
- 2: Draw v from $\operatorname{Unif}(\mathbb{S}_x^{d-2})$.
- 3: Draw θ from Unif $(0, 2\pi)$.
- 4: Set $\theta_{\min} = \theta 2\pi$ and $\theta_{\max} = \theta$.
- 5: repeat:
- 6: Draw θ from Unif $(\theta_{\min}, \theta_{\max})$.
- 7: Set $y = \cos(\theta)x + \sin(\theta)v$.
- 8: $\mathbf{if}\theta < \mathbf{0}$
- 9: $\theta_{\min} = \theta$
- 10: else
- 11: $\theta_{\max} = \theta$
- 12: until p(y) > t



- 1: Draw t from Unif(0, p(x)).
- 2: Draw v from $\text{Unif}(\mathbb{S}_x^{d-2})$.
- 3: Draw θ from Unif $(0, 2\pi)$.
- 4: Set $\theta_{\min} = \theta 2\pi$ and $\theta_{\max} = \theta$.
- 5: repeat:
- 6: Draw θ from Unif $(\theta_{\min}, \theta_{\max})$.
- 7: Set $y = \cos(\theta)x + \sin(\theta)v$.
- 8: $\mathbf{if}\theta < \mathbf{0}$
- 9: $\theta_{\min} = \theta$
- 10: else
- 11: $\theta_{\max} = \theta$
- 12: until p(y) > t



- 1: Draw t from Unif(0, p(x)).
- 2: Draw v from $\operatorname{Unif}(\mathbb{S}^{d-2}_{x})$.
- 3: Draw θ from Unif $(0, 2\pi)$.
- 4: Set $\theta_{\min} = \theta 2\pi$ and $\theta_{\max} = \theta$.
- 5: repeat:
- 6: Draw θ from Unif $(\theta_{\min}, \theta_{\max})$.
- 7: Set $y = \cos(\theta)x + \sin(\theta)v$.
- 8: $\mathbf{if}\theta < \mathbf{0}$
- 9: $\theta_{\min} = \theta$
- 10: else
- 11: $heta_{\max} = heta$
- 12: **until** p(y) > t



- 1: Draw t from Unif(0, p(x)). 2: Draw v from Unif (\mathbb{S}^{d-2}) . 3: Draw θ from Unif $(0, 2\pi)$. 4: Set $\theta_{\min} = \theta - 2\pi$ and $\theta_{\max} = \theta$. 5: repeat: Draw θ from Unif $(\theta_{\min}, \theta_{\max})$. 6: Set $y = \cos(\theta)x + \sin(\theta)y$. 7: $if\theta < 0$ 8: 9: $\theta_{\min} = \theta$ else 10: $\theta_{max} = \theta$ 11:
- 12: **until** p(y) > t



1: Draw t from Unif(0, p(x)). 2: Draw v from Unif (\mathbb{S}^{d-2}) . 3: Draw θ from Unif $(0, 2\pi)$. 4: Set $\theta_{\min} = \theta - 2\pi$ and $\theta_{\max} = \theta$. 5: repeat: Draw θ from Unif $(\theta_{\min}, \theta_{\max})$. 6: Set $y = \cos(\theta)x + \sin(\theta)y$. 7: $if\theta < 0$ 8: 9: $\theta_{\min} = \theta$ 10: else $\theta_{max} = \theta$ 11:12: **until** p(y) > t



- 1: Draw t from Unif(0, p(x)).
- 2: Draw v from $\operatorname{Unif}(\mathbb{S}_x^{d-2})$.
- 3: Draw θ from Unif $(0, 2\pi)$.
- 4: Set $\theta_{\min} = \theta 2\pi$ and $\theta_{\max} = \theta$.
- 5: repeat:
- 6: Draw θ from Unif $(\theta_{\min}, \theta_{\max})$.
- 7: Set $y = \cos(\theta)x + \sin(\theta)v$.
- 8: if $\theta < 0$
- 9: $\theta_{\min} = \theta$
- 10: else
- 11: $\theta_{\max} = \theta$
- 12: **until** p(y) > t



- 1: Draw t from Unif (0, p(x)). 2: Draw v from Unif (\mathbb{S}_x^{d-2}) . 3: Draw θ from Unif $(0, 2\pi)$. 4: Set $\theta_{\min} = \theta - 2\pi$ and $\theta_{\max} = \theta$. 5: repeat: 6: Draw θ from Unif $(\theta_{\min}, \theta_{\max})$. 7: Set $y = \cos(\theta)x + \sin(\theta)v$. 8: if $\theta < 0$ 9: $\theta_{\min} = \theta$
- 10: else

11:
$$heta_{\max} = heta$$

12: **until** p(y) > t



- 1: Draw t from Unif (0, p(x)). 2: Draw v from Unif (\mathbb{S}_x^{d-2}) . 3: Draw θ from Unif $(0, 2\pi)$. 4: Set $\theta_{\min} = \theta - 2\pi$ and $\theta_{\max} = \theta$. 5: repeat: 6: Draw θ from Unif $(\theta_{\min}, \theta_{\max})$. 7: Set $y = \cos(\theta)x + \sin(\theta)v$. 8: if $\theta < 0$ 9: $\theta_{\min} = \theta$
- 10: else

11:
$$\theta_{\max} = \theta$$

12: **until** p(y) > t



- 1: Draw t from Unif(0, p(x)). 2: Draw v from Unif (\mathbb{S}^{d-2}) . 3: Draw θ from Unif $(0, 2\pi)$. 4: Set $\theta_{\min} = \theta - 2\pi$ and $\theta_{\max} = \theta$. 5: repeat: Draw θ from Unif $(\theta_{\min}, \theta_{\max})$. 6: Set $y = \cos(\theta)x + \sin(\theta)y$. 7: $\mathbf{i}\mathbf{f}\theta < 0$ 8: 9: $\theta_{\min} = \theta$ 10: else
- 11: $heta_{\max} = heta$
- 12: **until** p(y) > t



- 1: Draw t from Unif(0, p(x)). 2: Draw v from Unif (\mathbb{S}^{d-2}) . 3: Draw θ from Unif $(0, 2\pi)$. 4: Set $\theta_{\min} = \theta - 2\pi$ and $\theta_{\max} = \theta$. 5: repeat: Draw θ from Unif $(\theta_{\min}, \theta_{\max})$. 6: Set $y = \cos(\theta)x + \sin(\theta)y$. 7: $\mathbf{i}\mathbf{f}\theta < 0$ 8: 9: $\theta_{\min} = \theta$ 10: else
- 11: $heta_{\mathsf{max}} = heta$
- 12: **until** p(y) > t



- 1: Draw t from Unif (0, p(x)). 2: Draw v from Unif (\mathbb{S}_x^{d-2}) . 3: Draw θ from Unif $(0, 2\pi)$. 4: Set $\theta_{\min} = \theta - 2\pi$ and $\theta_{\max} = \theta$. 5: repeat: 6: Draw θ from Unif $(\theta_{\min}, \theta_{\max})$. 7: Set $y = \cos(\theta)x + \sin(\theta)v$. 8: if $\theta < 0$ 9: $\theta_{\min} = \theta$
- 10: else

11:
$$heta_{\max} = heta$$

12: **until** p(y) > t



1: Draw t from Unif(0, p(x)). 2: Draw v from Unif (\mathbb{S}^{d-2}) . 3: Draw θ from Unif $(0, 2\pi)$. 4: Set $\theta_{\min} = \theta - 2\pi$ and $\theta_{\max} = \theta$. 5: repeat: Draw θ from Unif $(\theta_{\min}, \theta_{\max})$. 6: Set $y = \cos(\theta)x + \sin(\theta)y$. 7: $if\theta < 0$ 8: 9: $\theta_{\min} = \theta$ else 10: $\theta_{\max} = \theta$ 11: 12: **until** p(y) > t



Validity of the geodesic slice sampler on the sphere

Call the resulting kernel H.

Theorem (Habeck, H., Kodgirwar, and Rudolf 2023) The kernel H is reversible w.r.t. π . Moreover, assume

$$\alpha \coloneqq \sup_{x \in \mathbb{S}^{d-1}} p(x) < \infty.$$

Then for
$$\rho = 1 - \frac{\beta}{2\pi \ \alpha} \frac{\operatorname{vol}(C)}{\operatorname{vol}(\mathbb{S}^{d-2})} \in (0,1)$$
 holds
$$\sup_{x \in \mathbb{S}^{d-1}} \|H^n(x,\cdot) - \pi\|_{tv} \le \rho^n, \qquad \forall n \in \mathbb{N},$$

for some $C \in \mathcal{B}(\mathbb{S}^{d-1})$ with $\operatorname{vol}(C) > 0$ and $\beta \coloneqq \inf_{x \in C} p(x) > 0$.
Geodesics

Definition

A curve $\gamma : \mathbb{R} \supseteq I \to M$ is called *geodesic* if the covariant derivative of the velocity vector field is zero everywhere.

▶ The great circles are the geodesics of the sphere.

We call a Riemannian manifold *geodesically complete* if for all $x \in M$, $v \in T_xM$ there exists a unique geodesic

$$\gamma_{(x,v)}:\mathbb{R} o \mathsf{M}$$
 such that $\gamma_{(x,v)}(0)=x,\;rac{\mathrm{d}\gamma_{(x,v)}}{\mathrm{d}t}|_0=v.$

Choosing a random geodesic

We can turn the unit tangent spheres of a Riemannian manifold (M, \mathfrak{g})

$$U_x\mathsf{M}:=\{v\in T_x\mathsf{M}\mid \mathfrak{g}_x(v,v)=1\}$$

into Riemannian manifolds isometric to \mathbb{S}^{d-1} .

Let $\nu_{\mathfrak{q},x}$ the corresponding Riemannian measure and define the probability measure on U_xM

$$\mu_{\mathsf{x}} := \operatorname{vol}(\mathbb{S}^{d-1})^{-1}\nu_{\mathfrak{g},\mathsf{x}}.$$

We can sample µ_x using any isometry between the inner product spaces (T_xM, g_x) and (ℝ^d, ⟨·, ·⟩).

An additional complication

Problem: The geodesics of a Riemannian manifold are in general not closed.



Determine "approximation" of the intersection of the level set and the geodesic with techniques³ from 1-dim. slice sampling.

³Neal 2003.

- 1: Draw t from Unif(0, p(x)).
- 2: Draw v from μ_x .



Parameters: $m \in \mathbb{N}$, w > 0Input: $x \in M$

- 1: Draw t from Unif(0, p(x)).
- 2: Draw v from μ_x .



Parameters: $m \in \mathbb{N}$, w > 0Input: $x \in M$

- 1: Draw t from Unif(0, p(x)).
- 2: Draw v from μ_x .



Parameters: $m \in \mathbb{N}$, w > 0Input: $x \in M$

1: Draw t from Unif(0, p(x)).

2: Draw v from μ_x .



Output: $\gamma_{(x,v)}(\theta) \in M$

Mareike Hasenpflug (Uni. Passau) Geodesic slice sampling on Riemannian manifolds

Parameters: $m \in \mathbb{N}$, w > 0Input: $x \in M$

1: Draw t from Unif(0, p(x)).

2: Draw v from μ_x .



Parameters: $m \in \mathbb{N}$, w > 0Input: $x \in M$

- 1: Draw t from Unif(0, p(x)).
- 2: Draw v from μ_x .



Parameters: $m \in \mathbb{N}$, w > 0Input: $x \in M$

1: Draw t from Unif(0, p(x)).

2: Draw v from μ_x .

level set to level t





Output: $\gamma_{(x,y)}(\theta) \in M$

 $\{1, 2, 3, 4\}$



Output: $\gamma_{(x,y)}(\theta) \in M$

 $\{1, 2, 3, 4\}$



Output: $\gamma_{(x,y)}(\theta) \in M$

 $m = 4, \iota = 1$









13: Draw θ_h from Unif ((0, r - l)). 14: Set $\theta := \theta_h - \mathbb{1}_{\{\theta_h > r\}}(r-l)$. 15: Set $\theta_{\min} := \theta_h$ and $\theta_{\max} := \theta_h$. 16: while $(\gamma_{(x,y)}(\theta)) \leq t$ do if $\theta_h \in [\theta_{\min}, r - I]$ then 17: Set $\theta_{\min} = \theta_{h}$. 18: else 19. Set $\theta_{max} = \theta_h$. 20: Draw θ_h from 21: Unif $((0, \theta_{\max}) \cup [\theta_{\min}, r-I))$. Set $\theta = \theta_h - \mathbb{1}_{\{\theta_h > r\}}(r-l)$. 22: **Output:** $\gamma_{(x,y)}(\theta) \in M$



13: Draw θ_h from Unif ((0, r - l)). 14: Set $\theta := \theta_h - \mathbb{1}_{\{\theta_h > r\}}(r-l)$. 15: Set $\theta_{\min} := \theta_h$ and $\theta_{\max} := \theta_h$. 16: while $(\gamma_{(x,y)}(\theta)) \leq t$ do if $\theta_h \in [\theta_{\min}, r-l)$ then 17: Set $\theta_{\min} = \theta_{h}$. 18: else 19. Set $\theta_{max} = \theta_h$. 20: Draw θ_h from 21: Unif $((0, \theta_{\max}) \cup [\theta_{\min}, r-l))$. Set $\theta = \theta_h - \mathbb{1}_{\{\theta_h > r\}}(r-l)$. 22: **Output:** $\gamma_{(x,y)}(\theta) \in M$



13: Draw θ_h from Unif ((0, r - l)). 14: Set $\theta := \theta_h - \mathbb{1}_{\{\theta_h > r\}}(r-l)$. 15: Set $\theta_{\min} := \theta_h$ and $\theta_{\max} := \theta_h$. 16: while $(\gamma_{(x,y)}(\theta)) \leq t$ do if $\theta_h \in [\theta_{\min}, r-l]$ then 17: Set $\theta_{\min} = \theta_{h}$. 18: else 19. Set $\theta_{max} = \theta_h$. 20: Draw θ_h from 21: Unif $((0, \theta_{\max}) \cup [\theta_{\min}, r-l))$. Set $\theta = \theta_h - \mathbb{1}_{\{\theta_h > r\}}(r-l)$. 22:



Input: $x \in M$

13: Draw θ_h from Unif ((0, r - l)). 14: Set $\theta := \theta_h - \mathbb{1}_{\{\theta_h > r\}}(r-l)$. 15: Set $\theta_{\min} := \theta_h$ and $\theta_{\max} := \theta_h$. 16: while $(\gamma_{(x,y)}(\theta)) \leq t$ do if $\theta_h \in [\theta_{\min}, r-l)$ then 17: Set $\theta_{\min} = \theta_{h}$. 18: else 19. Set $\theta_{max} = \theta_h$. 20: Draw θ_h from 21: Unif $((0, \theta_{\max}) \cup [\theta_{\min}, r-l))$. Set $\theta = \theta_h - \mathbb{1}_{\{\theta_h > r\}}(r-l)$. 22:



Parameters: $m \in \mathbb{N}$, w > 0Input: $x \in M$

13: Draw θ_h from Unif ((0, r - l)). 14: Set $\theta := \theta_h - \mathbb{1}_{\{\theta_h > r\}}(r - I)$. 15: Set $\theta_{\min} := \theta_h$ and $\theta_{\max} := \theta_h$. 16: while $(\gamma_{(x,y)}(\theta)) \leq t$ do if $\theta_h \in [\theta_{\min}, r-l]$ then 17: Set $\theta_{\min} = \theta_{h}$. 18: else 19: Set $\theta_{max} = \theta_h$. 20: Draw θ_h from 21: Unif $((0, \theta_{\max}) \cup [\theta_{\min}, r - I))$. Set $\theta = \theta_h - \mathbb{1}_{\{\theta_h > r\}}(r-l)$. 22:



Stepping-out and shrinkage procedure

▶ Denote by $L(x, v, t) := \{\theta \in \mathbb{R} \mid p(\gamma_{(x,v)}(\theta)) > t\}$ intersection of L(t) and $\gamma_{(x,v)}$.

⁴H., Natarovskii, and Rudolf 2023.

Stepping-out and shrinkage procedure

- ▶ Denote by $L(x, v, t) := \{\theta \in \mathbb{R} \mid p(\gamma_{(x,v)}(\theta)) > t\}$ intersection of L(t) and $\gamma_{(x,v)}$.
- Define random variables

 $\Upsilon \sim \operatorname{Unif}(0, w)$ $L_i := -\Upsilon - w(i-1)$ $R_j := -\Upsilon + jw,$ $i, j \in \mathbb{N},$

stopping times, where $J \sim \text{Unif}(\{1, \ldots, m\})$,

$$\tau_{L(x,v,t)} := \min\{i \in \mathbb{N} \mid L_i \notin L(x,v,t)\} \land J,$$

$$\mathfrak{T}_{L(x,v,t)} := \min\{j \in \mathbb{N} \mid R_j \notin L(x,v,t)\} \land (m+1-J)\}$$

distribution on \mathbb{R}^2 of the stepping-out procedure approximating L(x, v, t) $\xi_{L(x,v,t)} := \operatorname{Law} \left(L_{\tau_{L(x,v,t)}}, R_{\mathfrak{T}_{L(x,v,t)}} \right).$

Mareike Hasenpflug (Uni. Passau) Geodesic slice sampling on Riemannian manifolds

⁴H., Natarovskii, and Rudolf 2023.

Stepping-out and shrinkage procedure

- ▶ Denote by $L(x, v, t) := \{\theta \in \mathbb{R} \mid p(\gamma_{(x,v)}(\theta)) > t\}$ intersection of L(t) and $\gamma_{(x,v)}$.
- Define random variables

 $\Upsilon \sim \operatorname{Unif}(0, w)$ $L_i := -\Upsilon - w(i-1)$ $R_j := -\Upsilon + jw,$ $i, j \in \mathbb{N},$

stopping times, where $J \sim \text{Unif}(\{1, \ldots, m\})$,

$$\tau_{\mathcal{L}(x,v,t)} := \min\{i \in \mathbb{N} \mid L_i \notin \mathcal{L}(x,v,t)\} \land J,$$

$$\mathfrak{T}_{\mathcal{L}(x,v,t)} := \min\{j \in \mathbb{N} \mid R_j \notin \mathcal{L}(x,v,t)\} \land (m+1-J).$$

distribution on \mathbb{R}^2 of the stepping-out procedure approximating L(x, v, t) $\xi_{L(x,v,t)} := \operatorname{Law}\left(L_{\tau_{L(x,v,t)}}, R_{\mathfrak{T}_{L(x,v,t)}}\right).$

► Let $Q_{L(x,v,t)}^{\ell,r}$ the distribution⁴ of the shrinkage procedure drawing from $L(x, v, t) \cap (\ell, r)$.

⁴H., Natarovskii, and Rudolf 2023.

Validity of the geodesic slice sampler

Kernel of the geodesic slice sampler for $x \in M$, $A \in \mathcal{B}(M)$:

$$\begin{aligned} \mathcal{K}(x,\mathsf{A}) &= \frac{1}{p(x)} \int_{0}^{p(x)} \int_{U_{x}\mathsf{M}} \int_{\mathbb{R}^{2}} \int_{L(x,v,t)\cap(\ell,r)} \mathbb{1}_{\mathsf{A}} \left(\gamma_{(x,v)}(\theta) \right) \\ & Q_{L(x,v,t)}^{\ell,r}(\mathrm{d}\theta) \, \xi_{L(x,v,t)}(\mathrm{d}(\ell,r)) \, \mu_{x}(\mathrm{d}v) \, \mathrm{d}t \end{aligned}$$

Theorem (Durmus, Gruffaz, H., and Rudolf 2023)

The kernel K is reversible with respect to π .

Numerical experiments: von Mises-Fisher distribution

Stiefel manifold:

$$\operatorname{Stiefel}(d,k) := \{ \Gamma \in \mathbb{R}^{d \times k} \mid \Gamma^{\top} \Gamma = \operatorname{Id}_k \}$$

Unnormalised density of the von Mises-Fisher distribution with parameter $m{F} \in \mathbb{R}^{d imes k}$

$$p_{\mathrm{vMF}}(\Gamma) = \exp\left(\mathrm{Tr}(F^{\top}\Gamma)
ight), \qquad \Gamma \in \mathrm{Stiefel}(d,k)$$

Choose

$$d = 30, \ k = 2,$$
 and $F = \begin{pmatrix} 1 & 0 \\ 0 & \lambda \\ 0 & 0 \\ \vdots & \vdots \end{pmatrix}, \lambda \in \{1, 10, 100\}.$

Effective sample size of $\ln(p_{\rm vMF})$ averaged over 10 repetitions:

λ	1	10	100
$\begin{array}{l} {\rm GSS} w=5,m=1\\ {\rm RMH}\\ {\rm GeoRMH} \end{array}$	[28375, 34262, 37405]	[4901, 5283, 5477]	[1153, 1328, 1453]
	[50772 , 54243, 59906]	[1492,2314,3214]	[669,878,998]
	[49007, 57195, 68948]	[1978,2336,3217]	[682,870,1075]

Numerical experiments: Functional connectivity network of the brain (Mantoux et al. 2021)

Model functional connectivity network of the brain:

$$A_j = X_j \operatorname{diag}(a_j) X_j^\top + \varepsilon_j, \qquad j \in \{1, \dots, N\}$$

where

$$X_j \stackrel{iid}{\sim} \mathrm{vMF}(F), \quad a_j \stackrel{iid}{\sim} \mathcal{N}(\mu, \sigma_a^2 \mathrm{Id}_k), \quad \varepsilon_j \stackrel{iid}{\sim} \mathcal{N}(0, \sigma_{\varepsilon}^2 \mathrm{Id}_{(d+1)/2}).$$

- ▶ Use data from 812 subjects from Human Connectome Project⁵.
- ▶ Estimate the parameters $F, \mu, \sigma_a, \sigma_\varepsilon$ with 1000 MCMC-SAEM.
- ► Use either 20 geodesic slice sampling or 80 RMH iterations within MCMC-SAEM.

 $^{{}^{\}tt 5} https://www.humanconnectome.org/study/hcp-young-adult/document/extensively-processed-fmridata-documentation}$

Numerical experiments: von Mises-Fisher distribution



Mareike Hasenpflug (Uni. Passau) Geodesic slice sampling on Riemannian manifolds

References |

- Durmus, Gruffaz, H., and Rudolf (2023). Geodesic slice sampling on Riemannian manifolds. arXiv: 2312.00417.
- H., Natarovskii, and Rudolf (2023). Reversibility of elliptical slice sampling revisited. arXiv: 2301.02426.
- Habeck, H., Kodgirwar, and Rudolf (2023). *Geodesic slice sampling on the sphere*. URL: https://arxiv.org/abs/2301.08056.
- Mantoux, Couvy-Duchesne, Cacciamani, Epelbaum, Durrleman, and Allassonnière (2021). "Understanding the variability in graph data sets through statistical modeling on the Stiefel manifold". In: Entropy 23.4.
- Natarovskii, Rudolf, and Sprungk (2021). "Quantitative spectral gap estimate and Wasserstein contraction of simple slice sampling". In: The Annals of Applied Probability 31.2, pp. 806–825.
- Neal (2003). "Slice sampling". In: The Annals of Statistics 31.3, pp. 705–767.
- Rudolf and Schär (2024). "Dimension-independent spectral gap of polar slice sampling". In: Statistics and computing 34, 20.

Schär (2023). "Wasserstein contraction and spectral gap of slice sampling revisited". In: *Electronic Journal of Probability* 28, pp. 1–28.

Distributions on manifolds

We can incorporate constraints or dependencies directly into a model via manifold state spaces.

Example (Mantoux et al. 2021) Model adjacency matrix $A \in \mathbb{R}^{d \times d}$ of an undirected graph through its eigendecomposition: $A = X \operatorname{diag}(a) X^{\top} + \operatorname{noise}, \qquad a \in \mathbb{R}^k, X \in \operatorname{Stiefel}(d, k).$

 Applying e.g. Bayesian inference or MCMC-SAEM requires sampling from distributions on manifolds.

Some differential geometry

Definition

We call coordinate neighbourhoods (U, φ), (V, ψ) C^{∞} -compatible if $\psi \circ \varphi^{-1}$ and $\varphi \circ \psi^{-1}$ are infinitely often differentiable.

A manifold M with a maximal collection of C^{∞} -compatible coordinate neighbourhoods that cover M is called *smooth*.



Some differential geometry

Definition

We call coordinate neighbourhoods (U, φ), (V, ψ) C^{∞} -compatible if $\psi \circ \varphi^{-1}$ and $\varphi \circ \psi^{-1}$ are infinitely often differentiable.

A manifold M with a maximal collection of C^{∞} -compatible coordinate neighbourhoods that cover M is called *smooth*.



Some differential geometry

Definition

We call coordinate neighbourhoods (U, φ), (V, ψ) C^{∞} -compatible if $\psi \circ \varphi^{-1}$ and $\varphi \circ \psi^{-1}$ are infinitely often differentiable.

A manifold M with a maximal collection of C^{∞} -compatible coordinate neighbourhoods that cover M is called *smooth*.


Definition

A function $f : A \to \mathbb{R}$ on an open set $A \subseteq M$ is called C^{∞} -function if $f \circ \varphi^{-1}$ is infinitely often differentiable for all coordinate neighbourhoods (U, φ) . Set $C_{M}^{\infty}(x) := \{f : A \to \mathbb{R}C^{\infty}$ -function $| x \in A \}$.



Definition

A function $f : A \to \mathbb{R}$ on an open set $A \subseteq M$ is called C^{∞} -function if $f \circ \varphi^{-1}$ is infinitely often differentiable for all coordinate neighbourhoods (U, φ) . Set $C^{\infty}_{M}(x) := \{f : A \to \mathbb{R}C^{\infty}$ -function $| x \in A \}$.



Definition

An operator $X: C^\infty_\mathsf{M}(x) o \mathbb{R}$ is called *tangent vector* to M at $x \in \mathsf{M}$ if

$$X(\alpha f + \beta g) = \alpha X(f) + \beta X(g),$$

$$X(fg) = X(f)g(x) + f(x)X(g)$$

for all $\alpha, \beta \in \mathbb{R}$, $f, g \in C^{\infty}_{\mathsf{M}}(x)$. The \mathbb{R} -vector space $T_x\mathsf{M} := \{X \text{ tangent vector to } \mathsf{M} \text{ at } x\}$ is called *tangent space*.



Definition

A mapping $\mathfrak{g} : M \ni x \mapsto \mathfrak{g}_x$ is called *Riemannian metric* if \mathfrak{g}_x is an inner product on T_xM for all $x \in M$ and some smoothness condition is satisfied.

A smooth manifold M with a Riemannian metric is called *Riemannian*.



Definition

A mapping $\mathfrak{g} : M \ni x \mapsto \mathfrak{g}_x$ is called *Riemannian metric* if \mathfrak{g}_x is an inner product on T_xM for all $x \in M$ and some smoothness condition is satisfied.

A smooth manifold M with a Riemannian metric is called *Riemannian*.



Let M be a Riemannian manifold with Borel- σ -algebra $\mathcal{B}(M)$, and

- (U, φ) coordinate neighbourhood,
- √det(g, φ)(x) square root of determinant of Gram matrix of coordinate frames associated to (U, φ) with respect to g_x.

Define the *Riemannian measure* as

$$\nu_{\mathfrak{g}}(\mathsf{A}) := \int_{\varphi \ (\mathsf{U} \)} \Big(\quad \mathbb{1}_{\mathsf{A}} \cdot \sqrt{\mathsf{det}(\mathfrak{g}, \varphi \)} \Big) \circ \varphi^{-1}(z) \ \mathrm{d} z, \qquad \mathsf{A} \subseteq \mathsf{U}, \mathsf{A} \in \mathcal{B}(\mathsf{M}).$$

Let M be a Riemannian manifold with Borel- σ -algebra $\mathcal{B}(M)$, and

- $\{(U_i, \varphi_i)\}_{i \in \mathbb{N}}$ countable atlas of M,
- $\sqrt{\det(\mathfrak{g},\varphi_i)}(x)$ square root of determinant of Gram matrix of coordinate frames associated to (U_i,φ_i) with respect to \mathfrak{g}_x ,
- $\{\rho_i\}_{i\in\mathbb{N}}$ partition of unity subordinate to $\{U_i\}_{i\in\mathbb{N}}$, (i.a. $\operatorname{supp}\rho_i \subseteq U_i$, $\sum_{i\in\mathbb{N}}\rho_i \equiv 1$).

Define the *Riemannian measure* as

$$\nu_{\mathfrak{g}}(\mathsf{A}) := \sum_{i \in \mathbb{N}} \int_{\varphi_i(\mathsf{U}_i)} \left(\rho_i \cdot \mathbb{1}_{\mathsf{A}} \cdot \sqrt{\mathsf{det}(\mathfrak{g}, \varphi_i)} \right) \, \circ \, \varphi_i^{-1}(z) \, \mathrm{d}z, \qquad \mathsf{A} \in \mathcal{B}(\mathsf{M}).$$