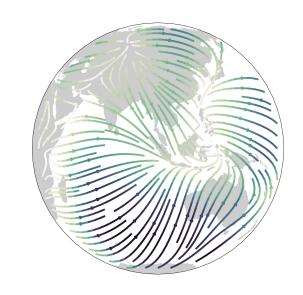
Riemannian Continuous Normalizing Flows

Learning expressive probability distributions on manifolds

Emile Mathieu and Maximilian Nickel



FACEBOOK AI



Overview:

Extend continuous normalizing flows to manifolds

- Flows on manifolds as solution of ODEs
- Continuous change of variables for manifold-valued variables
- Assess model expressiveness on earth science data

Motivation

$$\psi: \mathbb{R}^d o \mathcal{M}$$
 $P_{ heta} = (\psi \circ \phi_{ heta})_{\sharp} P_0$ $\phi_{ heta}$ P_0 Normalizing Flow $\phi_{ heta \sharp} P_0$

Gemici et al, 2016

Ordinary Differential Equation & Flow

ODE
$$\dfrac{dz(t)}{dt}=f_{ heta}(z(t),t)$$
 vector field $f_{ heta}(z_1(0),0)$ $z_1(T)=\phi(z_1(0),T)$ $z_1(t)=\phi(z_1(0),t)$ $z_2(0)$

Flow
$$\phi(z):=\phi(z,T)$$
 $\phi:\mathcal{M}\to\mathcal{M}$ \mathcal{C}^1 - diffeomorphism if f_{θ} is \mathcal{C}^1 and bounded

Model
$$P_{ heta} = \phi_{\sharp} P_0$$

Likelihood

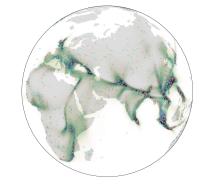
Continuous change of variable

$$\frac{\partial \log p_{\theta}(\boldsymbol{z}(t))}{\partial t} = -\operatorname{div}(\boldsymbol{f}_{\theta}(\boldsymbol{z}(t), t)) = -|G(\boldsymbol{z}(t))|^{-\frac{1}{2}} \operatorname{tr}\left(\frac{\partial \sqrt{|G(\boldsymbol{z}(t))|}\boldsymbol{f}_{\theta}(\boldsymbol{z}(t), t)}{\partial \boldsymbol{z}}\right)$$
(1)

with z some local coordinates, e.g. $z=(\theta,\phi)$ polar coordinates for \mathbb{S}^2

Stochastic estimator
$$\operatorname{div}(\boldsymbol{f}_{\theta}(\boldsymbol{z}(t),t)) = |G(\boldsymbol{z}(t))|^{-\frac{1}{2}} \mathbb{E}_{p(\boldsymbol{\epsilon})} \left| \boldsymbol{\epsilon}^{\mathsf{T}} \frac{\partial \sqrt{|G(\boldsymbol{z}(t))|} \boldsymbol{f}_{\theta}(\boldsymbol{z}(t),t)}{\partial \boldsymbol{z}} \boldsymbol{\epsilon} \right|$$

Experimental Results: Earth Science



	Volcano	Earthquake	Flood	Fire
Mixture vMF ■	$-0.31_{\pm 0.07}$	$0.59_{\pm 0.01}$	$1.09_{\pm 0.01}$	$-0.23_{\pm 0.02}$
Stereographic	$-0.64_{\pm 0.20}$	$0.43_{\pm 0.06}$	$0.99_{\pm 0.04}$	$-0.40_{\pm 0.06}$
Riemannian	$-0.97_{\pm 0.15}$	$0.18_{\pm 0.05}$	$0.90_{\pm 0.03}$	$-0.66_{\pm 0.05}$
Learning curves	3 2 1 0 -1 0 1000 2000 3000 epochs	2- 1- 0 500 1000 epochs	2- 1- 0 500 1000 epochs	2- 1- 0 500 1000 epochs
Data size	829	6124	4877	12810

Thank you for your attention!

Don't be shy and attend our poster session for a chat:)