Equivariant Manifold Flows

Katsman et al. 2021

Tommy He

Outline

Summary

Equivariance has proven to be a powerful inductive bias so extend it to continuous normalizing manifolds flows!

Provide a framework for explicitly **incorporating symmetries into flows** over arbitrary manifolds that and prove the flows universally approximate distributions on closed manifolds

Experiment with learning gauge invariant densities over $SU(n)$ for quantum field theory and asteroid impacts over S^2

Related Work & Why Not Take Quotient

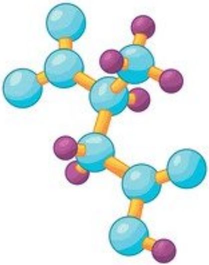
Boyda et al. 2020 first introduced equivariant flows only on $\mathbf{SU}(\mathbf{n})$, though used the **quotient** of the manifold by isometry subgroup and ensured equivariance with equivariant coupling layers under matrix conjugation, and could not generalize

Who knows what the quotient manifold looks like, in terms of its charts and boundaries, if it even exists? Even if doable, oddities can arise.

Examples

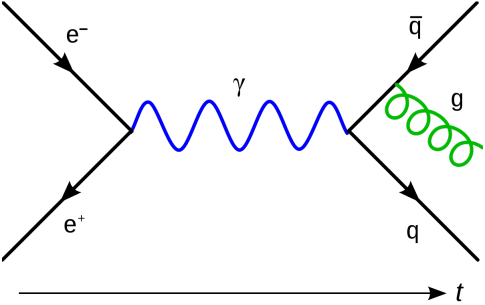
Asteroid Impacts on Earth: S^2

Protein Synthesis: $S^1 \times S^1$



Examples

Lattice Quantum Field Theory: $SU(n)$
Gauge Theory: $SU(n)$



Isometry Groups

Take some Riemmanian manifold $(\mathcal{M}, h)$ with tangent bundle $T\mathcal{M}$ and tangent spaces $T_p\mathcal{M} \forall p \in \mathcal{M}$.

Diffeomorphism $f : \mathcal{M} \to \mathcal{M}$ is a differentiable bijection with differentiable inverse

Isometry $f : \mathcal{M} \to \mathcal{M}$ is a diffeomorphism if it preserves distance

$$h(u, v) = h(D_p(f(u)), D_p(f(v))) \forall u, v \in T_p\mathcal{M}$$

The set of isometries forms the **isometry group** G where the action of $g \in G$ on $\mathcal{M}$ is denoted $L_g : \mathcal{M} \to \mathcal{M}, m \mapsto gm$

Normalizing Flows on Manifolds

Manifold Normalizing Flows is diffeomorphism $f : \mathcal{M} \to \mathcal{M}$ that transforms prior density to model density

$$f : \rho \mapsto \rho_f$$

Manifold Continuous Normalizing Flows is function $\gamma : [0, \infty) \to \mathcal{M}$ with base point z modeled by ODE

$$\frac{\mathrm{d}\gamma(t)}{\mathrm{d}t} = X(\gamma(t), t), \quad \gamma(0) = z$$

Let $F_{X,T} : \mathcal{M} \to \mathcal{M}$ map from base point z to $\gamma(t)$ under X.

Equivariance and Invariance Functions

Let $(\mathcal{N}, g)$ be some Riemmanian manifold. Then, $f : \mathcal{M} \to \mathcal{N}$ is **equivariant** if for all isometries $g_m : \mathcal{M} \to \mathcal{M}, g_n : \mathcal{N} \to \mathcal{N}$, the following commutes

$$\begin{array}{ccc} \mathcal{M} & \xrightarrow{g_m} & \mathcal{M} \\ \big\downarrow f & & \big\downarrow f \\ \mathcal{N} & \xrightarrow{g_n} & \mathcal{N} \end{array}$$

f is **invariant** if $g_m f = f$

Equivariant Vector Fields

Time-dependent vector field $X : \mathcal{M} \times [0, \infty) \to T\mathcal{M}$ is **equivariant** if the following commutes

$$\begin{tikzcd} \mathcal{M} \arrow[r, "L_g"] \arrow[d, "X"] & \mathcal{M} \arrow[d, "X"] \\ T\mathcal{M} \arrow[r, "DL_g"] & T\mathcal{M} \end{tikzcd}$$

Overview

Main theorems 1-3 to construct a G-invariant density from a G-invariant potential and show how easy it is to find a G-invariant potential.

$$
\begin{aligned}
& G\text{-invariant potential} \\
\xrightarrow{Thm.1}\ & G\text{-equivariant vector field} \\
\xrightarrow{Thm.2}\ & G\text{-equivariant flow} \\
\xrightarrow{Thm.3}\ & G\text{-invariant density}
\end{aligned}
$$

Thm. 1 G-invariant Potential $\Rightarrow$ G-equivariant Vector Field

$\Phi : \mathcal{M} \to \mathbb{R}$ a smooth G-invariant function (the potential), then the following commutes

$$\begin{array}{ccc}
\mathcal{M} & \xrightarrow{L_g} & \mathcal{M} \\
\Big\downarrow \nabla\Phi & & \Big\downarrow \nabla\Phi \\
T\mathcal{M} & \xrightarrow{DL_g} & T\mathcal{M}
\end{array}$$

$\nabla\Phi$ is G-equivariant vector field

Thm. 2 G-equivariant Vector Field $\Rightarrow$ G-equivariant Flow

Time-dependent vector field X on $\mathcal{M}$ with flow $F_{X,T}$

X is G-equivariant vector field $\iff$ $F_{X,T}$ is G-equivariant flow

Thm. 3 G-equivariant Flow $\Rightarrow$ G-equivariant Density

G-invariant density ρ (prior) with G-equivariant diffeomorphism f (flow)
$\Rightarrow$pushforward density ρ_f (target) is G-invariant

Expressivity

Can any equivariant flow be generated this way?

Expressivity

Can any equivariant flow be generated this way?

Yes! Mostly

Expressivity

Can any equivariant flow be generated this way?

Yes! Mostly

Any smooth invariant distribution over a closed (compact without boundaries) can be recovered up to KL

Thm. 4 Universal Expressivity

For smooth, non-vanishing target distribution π, distribution ρ_t from flow at time t, choose potential Φ as

$$\Phi(x) = \log\left(\frac{\pi(x)}{\rho_t(x)}\right)$$

If follows

$$\frac{\partial}{\partial x} D_{KL}(\rho_t \parallel \pi) = -\int_{\mathcal{M}} \pi \|\nabla g\|^2 dx$$

which means $\rho_t \to \pi$ in KL. Moreover, the diffeomorphism $\rho_t \mapsto \pi$ maps $x \mapsto \lim_{t\to\infty} u(t)$ for solution $u(t)$ to

$$\frac{\mathrm{d}u(t)}{\mathrm{d}t} = \nabla\Phi(t), u(0) = x$$

Construction of Φ

Invariance of Φ will leave some parameters unrestricted, so use a neural network to learn the distribution of the free parameters

Can be arbitrary due to freeness for equivariance, but we try to learn the best as it gets pushed through to the final density

Isotropy Invariance on S^2

Take S^2 with fixed point $(0, 0, 1)$. Then, the isometry group is the rotations through the z-axis around the xy-plane (stripes) with the z-axis as free

Take a 2-input neural network with z, t as inputs

Take a prior distribution that respects isometry invariance, in our case uniform

Conjugation Invariance on $SU(n)$

Note conjugation invariant implies it preserves eigenvalues, so the invariance is a permutation of eigenvalues

Take DeepSet, a permutation invariant neural network that acts on eigenvalues, and append time t

Take a prior distribution that respects matrix conjugation invariance, in our case Haar measure

Learning the Invariant Desntiy

(Thm. 1) Use automatic differentiation on Φ to get $\nabla\Phi$
(Thm. 2) Integrate step-wise over vector field to get flow as per *Riemmanian Continuous Normalizing Flows* (Mathieu & Nickel 2020)
(Thm. 3) Use the flow with training procedure

Training Paradigms

Learning **tractable sampler with exact distribution**

We don't have samples, so can't use negative log-likelihood $\rightarrow$ sample from the prior and compute KL divergence between probabilities

Learning **density with samples**

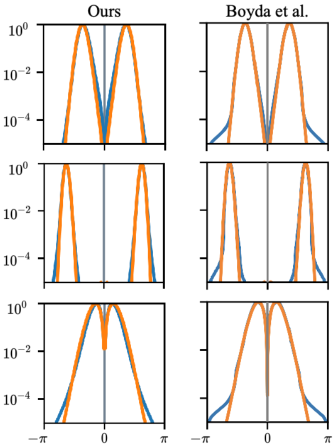
We sample the target density and flow to the prior and use negative log-likelihood

Experiments

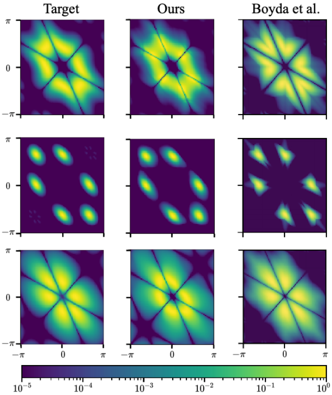
Learning $SU(n)$ **gauge equivariant flows**, specifically $SU(2)$ and $SU(3)$, and comparing with Boyda et al

Modeling **asteroid collision** and comparing with NMODE

$SU(n)$ gauge equivariant flows



$SU(2)$



$SU(3)$

Asteroid modeling

Isotropy invariant density in heatmap

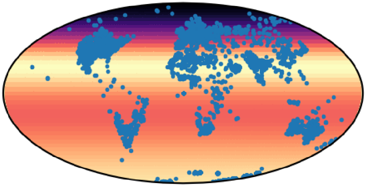
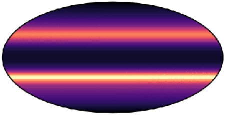
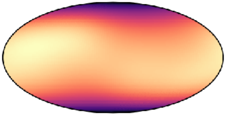


Figure: Comparison why NMODE & why equivariance matters

(a) Ground Truth



(b) Isotropy Equivariant Flow



(c) NMODE [29]

Concluding Remarks

Strengths

- First actual equivariant flow over generalized manifolds
- Introduces idea of deriving flows from potential functions

Weaknesses

- Numbers backing up experiments
- Experimentation not as tied to theoretical physics

Future Directions

- Construction of equivariant flows without going thorugh potential functions
- Generalization to open manifolds
- Extending construction of potential function

Thank you!

Questions?