Régression géodésique dans des groupes de Lie

Johan Aubray[1], Florence Nicol[1], Stéphane Puechmorel[1], Anne-Françoise Yao[2,3]

Université de Toulouse, Laboratoire ENAC, France, prenom.nom@enac.fr

Université Clermont Auvergne, Laboratoire de Mathématiques Blaise Pascal

CMAP, Ecole polytechnique, Palaiseau, France, anne.yao@uca.fr

Journées de Statistique et Optimisation en Occitanie
JS20 - Université de Perpignan
02-04 Avril 2025

Schedule

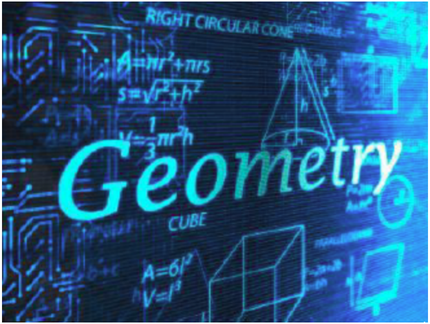
Motivations

Information Geometry

A multidisciplinary research domain

New interactions were discovered between several branches of science:

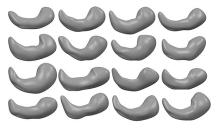
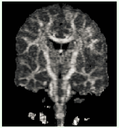
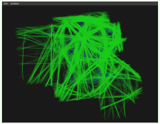
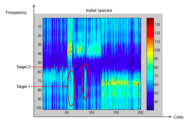
- Information Science (information theory, digital communications, statistical signal processing,),
- Mathematics (group theory, geometry and topology, probability, statistics etc.)
- Physics (geometric mechanics, thermodynamics, statistical physics, quantum mechanics etc.).

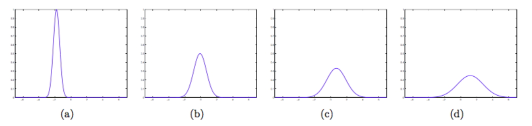


Non Euclidean data

In many applications, the data of interest are non-Euclidean

- curves, shapes, surfaces (computer vision, medical imaging, air transportation)
- signals, spectra (radar signal processing)
- SPD matrices, rotation matrices,...
- probability distributions



Information Geometry vs Geometric Statistics

Geometric Statistics (GS)

- Statistical analysis of **geometric objects living in differential manifolds** such as in computer vision and medical imaging analysis.
- Statistical analysis data lying in a known manifold:
  - directional data on the unit sphere $S\hat{}d$,
  - MRI signals of brain activity.

Information Geometry (IG)

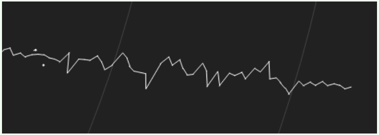
- It deals with **differential manifolds of statistical models** such as parameterized families of probability distributions;
- The space of parameters Θ is a smooth manifold, often an open set of $\mathbb{R}^n$, but may also be spheres, torus etc...
- Parameters θ are understood as local coordinates on the manifold Θ.

☞ GS and IG are linked by the same differential geometry framework **but the targeted geometric objects are different!**

Application to Air Transport Management (ATM)

Trajectory estimation

Estimating the true position of a mobile object (aircraft, drone, ...) from measured (therefore noisy) raw positions.

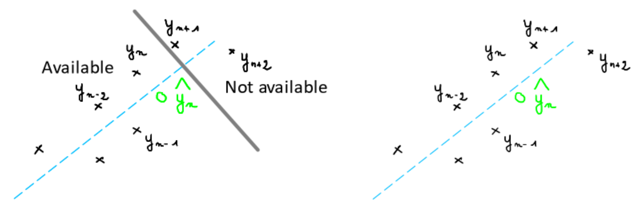


☞ This may be difficult with a highly maneuvering target.

Estimating the true position of a mobile object from noisy data

Two situations

- From past positions: **prediction**
  - Frequently solved with linear / extended / unscented Kalman filters, or particle filters
- From past and future positions: **smoothing**
  - Frequently solved with statistical regression techniques



☞ What comes next is based on **smoothing**

A geometrical representation

A new representation space

- Instead of considering the trajectory in the state space, the movement of the mobile object is considered as a sequence of displacements: rotations and translations in $\mathbb{R}^3$.
- This set is called the Special Euclidean group $SE(3)$.

A Lie group representation $SE(3)$

- a differentiable manifold
- a group in which both the product and inverse maps are smooth

☞ Switch to a geometrical point of view !
☞ The **Geometric Statistics** framework.

Tools of Differential Geometry

The differential geometry framework

Historical perspectives

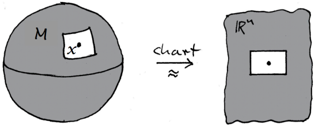
- **1827** “Disquisitiones generales circa superficies curvas" Carl Friedrich Gauss freely used local coordinates on a surface and he already had the idea of charts.
- **1854** “Über dir Hypothesen, welche der Geometrie zu Grunde liegen" Bernhard Riemann’s inaugural lecture laid the foundations of higher-dimensional differential geometry. He described manifolds that are direct translation of the German word “Mannigfaltigkeit”.
- **XIXth** century with the work of Henri Poincaré on homology
- **XXth** century with the modern definition of a manifold and transition functions.

Topological manifold

☞ Intuitively, a manifold is a generalization of curves and surfaces to higher dimensions.

Topological manifold

A (topological) **manifold** is a separable metrizable space M (Hausdorff space with a countable basis) which is locally Euclidian i.e. for all p in M, there exists an homeomorphism $\varphi : U \longrightarrow \varphi(U) \subset \mathbb{R}^n$, where U is an **open neighborhood** of p in M and $\varphi(U)$ **an open subset** in $\mathbb{R}^n$.



☞ A manifold locally looks like a Euclidian space $\mathbb{R}^n$ by using charts: every point has a neighborhood homeomorphic to an open subset of $\mathbb{R}^n$.
☞ Then, by gluing pieces of $\mathbb{R}^n$ together, we can build a more complicated object.

Charts and atlas

Chart and coordinate system

The pair (U, φ) is called a **chart** (or **local coordinates**) on M. The homeomorphism φ is called a **coordinate system**.

Dimension

The integer n is the **dimension** of M.

Atlas

A collection $\mathscr{A} = \{(U_\alpha, \varphi_\alpha), \alpha \in I\}$ of charts on M such that $\mathscr{A}$ covers M i.e.

$$M = \cup_{\alpha \in I} U_\alpha$$

is called an **atlas** on M.

Examples

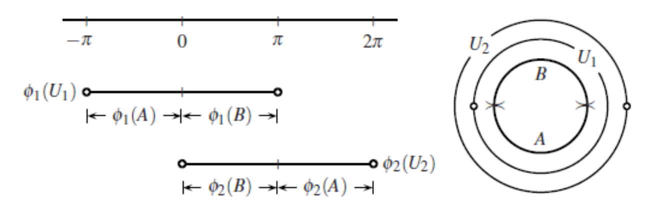
Euclidian space

The Euclidian space $\mathbb{R}^n$ with the trivial (global) chart $(\mathbb{R}^n, \mathbb{1}_{\mathbb{R}^n})$, where $\mathbb{1}_{\mathbb{R}^n}$ is the identity map.

Circle

The unit circle $\mathbb{S}^1 = \{e^{it} \in \mathbb{C}, 0 \leq t < 2\pi\}$, with two charts defined by

$$U_1 = \{e^{it} \in \mathbb{C} | -\pi < t < \pi\} \text{ and } \phi_1(e^{it}) = t, -\pi < t < \pi,$$
$$U_2 = \{e^{it} \in \mathbb{C} | 0 < t < 2\pi\} \text{ and } \phi_2(e^{it}) = t, 0 < t < 2\pi.$$



Examples

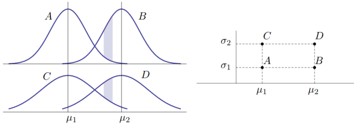
Matrix groups

- Special Orthogonal Group: $SO(n)$ rotations of $\mathbb{R}$
  - All matrices $R \in GL(n)$ such that $RR^T = I$ and $\det R = 1$.

Manifold of Gaussian distributions: a 2-dimensional manifold

$$M = \{p(x;\mu,\sigma), \mu \in \mathbb{R}, \sigma > 0\} \quad p(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

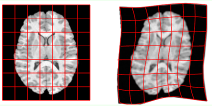
- Each probability distribution $p(x;\mu,\sigma)$ is viewed as one point in the manifold M.
- The mapping $\varphi(p(x;\mu,\sigma)) = (\mu,\sigma)$ is one possible coordinate system.
- The mapping $\psi(p(x;\mu,\sigma)) = (\mu,\sigma^2)$ is another coordinate system.



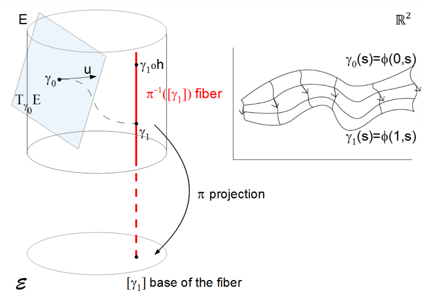
Examples

Shape spaces

- Space of curves: immersions $E = Imm([0,1],\mathbb{R}^d)$
- Space of diffeomorphisms: $Diff([0,1])$



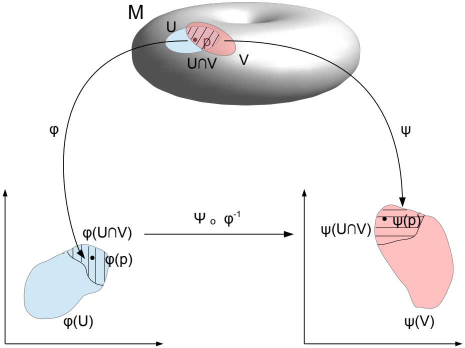
- Shape space: $Imm([0,1],\mathbb{R}^d)/Diff([0,1])$



Gluing charts together

Transition maps

The mappings $\varphi_{\alpha\beta} := \varphi_\alpha \circ \varphi_\beta^{-1} : \varphi_\beta(U_\alpha \cap U_\beta) \longrightarrow \varphi_\alpha(U_\alpha \cap U_\beta)$ are called **transition maps**.



- The transition maps correspond to **changes of charts** for the atlas $\mathscr{A}$.
- They play a crucial role for building manifolds and for defining their properties.

Role of transition maps

Properties on manifolds

- Properties of the manifold are encoded in the charts gluing: by changing properties on transition maps (smooth, differentiable, analytic, holomorphic, . . .), one may build different types of manifolds.
- Manifolds properties should be compatible with transition maps.

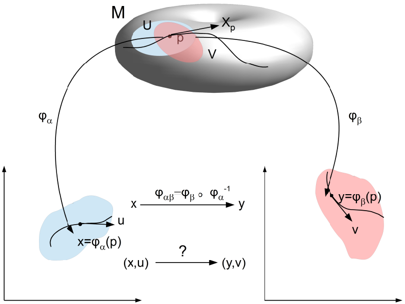
Differentiable manifolds

- If the charts are suitably compatible: **differentiable manifold**. One can then define the notion of tangent space and differentiable mapping.

Going off on a tangent (vector)

Basic ideas

- At any point p of a manifold M, one may define a tangent vector X_p as the derivative of a curve going through p.
- We glue two charts together by using transition maps: points $y = \varphi_{\alpha\beta}(x)$ as usual and velocity vectors $v = D_x \varphi_{\alpha\beta}(u)$ by using a linear operator.
- This notion has to be independent to the choice of charts.



Going off on a tangent (space)

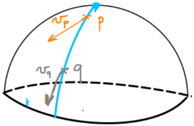
Tangent space T_pM

The set of all tangent vectors at p is a vector space called the tangent space at p, denoted T_pM

Tangent bundle TM

The union of all tangent spaces of M is the tangent bundle TM.

☞ Problem: how can we compare a vector $v_p \in T_p\mathscr{M}$ with a vector $v_q \in T_q\mathscr{M}$?



☞ Is is possible to generalize the notion of directional derivative of a function to vector fields ? **NO since *TM* is not a vector space!**

Connecting tangent spaces

A new structure on the manifold : the connexion ∇

An affine connection ∇ maps a couple of vector fields (X, Y) to a vector field $\nabla_X Y$ and satisfies the following properties:

1. $\nabla_X f = X(f)$,
2. $\nabla_{fX+Z} Y = f\nabla_X Y + \nabla_Z Y$,
3. $\nabla_X (fY + Z) = X(f)Y + f\nabla_X Y + \nabla_X Z$,

where $f : M \to \mathbb{R}$ is a smooth function. Remark that the last property corresponds to the Leibnitz rule

$$\nabla_X (fY + Z) = (\nabla_X f)Y + f\nabla_X Y + \nabla_X Z$$

- A new structure $(M, \mathscr{A}, \nabla)$.

Manifolds and geodesics

Geodesics as straight lines

- A geodesic in a manifold is the equivalent of a straight line.
- A curve γ is a geodesic if it has a "constant" velocity vector.

Geodesics as parallel-transport

A curve γ is called a geodesic if

$$\nabla_{\gamma'}\gamma' = \frac{D}{dt}\gamma'(t) = 0$$

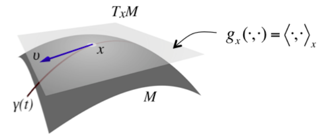
(we "parallel-transport" the $\gamma'(t)$ vector along the curve γ)

Riemannian metric

Riemannian manifold: $(M,\mathscr{A},g)$

- One can compute angles, lengths, distances using a **Riemannian metric** = scalar product locally defined on each tangent space

$$g_x(v,w) = v^T G(x) w, \quad x \in M, \quad v,w \in T_x M$$



The metric matrix

- We can represent the metric by a symmetric positive definite matrix

$$G(x) = (g_{ij}(x))_{i,j=1,\dots,n}$$

where $g_{ij}(x) = g_x(\partial_i, \partial_j) = <\partial_i, \partial_j>_x$.

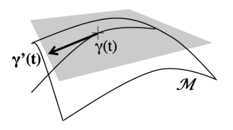
- The metric matrix $G(x) = (g_{ij}(x))_{i,j=1,\dots,n}$ is the **metric representation in local coordinates**.

Geodesics

Geodesic distance

- Length of a smooth path $\gamma : [0,1] \to M$ path in M

$$L(\gamma) = \int_0^1 \sqrt{g(\gamma'(t), \gamma'(t))}\mathrm{d}t.$$



- Paths of (locally) minimal length are called **"minimizing geodesics"** or "geodesics". The **geodesic distance** between two points is

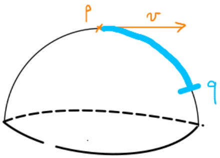
$$d(x,y) = \inf_{\gamma : \gamma(0)=x, \gamma(1)=y} L(\gamma).$$

☞ Geodesics correspond to straight lines in a manifold i.e. such that $\nabla_{\gamma'}\gamma' = 0$.
☞ Minimizing geodesics corresponds to paths of minimal length.
☞ These two notions coincide when the Riemannian structure is endowed with the Levi Civita connection $(M, \mathscr{A}, g, \nabla^{LC})$ (torsion free and metric compatible).

Exponential and logarithm on a manifold

- An initial point $p = \gamma(0)$ and an initial vector $v = \gamma'(0)$ are enough to define (at least locally) a unique geodesic
- This defines a map from the tangent bundle of $\mathcal{M}$ $T\mathcal{M} = \cup_p T_p\mathcal{M}$ to the manifold $\mathcal{M}$ called Exp:

$$\begin{aligned}\mathrm{Exp}: T\mathcal{M} &\to \mathcal{M} \\ (p, v) &\mapsto q\end{aligned}$$

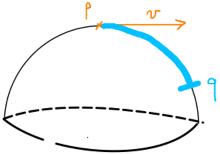


- This is the equivalent of $q = p + \vec{v}$ in $\mathbb{R}^n$

Exponential and logarithm on a manifold

- Likewise, one may define the inverse map called Log:

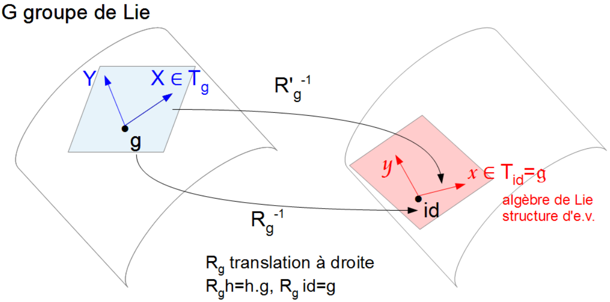
$$\mathrm{Log}_p : \mathcal{M} \to T_p\mathcal{M}$$
$$q \mapsto v$$



- This is the equivalent of $\vec{v} = \vec{pq} = q - p$ in $\mathbb{R}^n$

How about Lie groups?

- The group structure of a Lie group *G* makes it possible to bring most operations to the identity *e*
- The tangent space at *e* of *G* is called the Lie algebra of *G*, denoted $\mathfrak{g}$
  - it "linearises" the group and captures most of its properties



$\ll X, Y \gg_g$??? → $< R'^{-1}_g X, R'^{-1}_g Y >_{\mathfrak{g}}$

$||g'||^2 = \ll g', g' \gg_g$

Lie group: SE(3), Lie algebra: $\mathfrak{se}(3)$

- An element of $SE(3)$ can be represented by a 4-by-4 matrix

$$\begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix}, R \in SO(3), T \in \mathbb{R}^3:$$

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ 1 \end{pmatrix} = \begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \\ z_0 \\ 1 \end{pmatrix} = \begin{pmatrix} R \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + T \\ 1 \end{pmatrix}$$

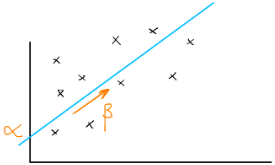
- An element of $\mathfrak{se}(3)$ can be represented by a 4-by-4 matrix:

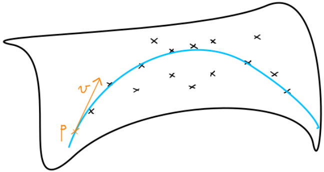
$$\begin{pmatrix} \Omega & U \\ 0 & 0 \end{pmatrix}, \Omega \text{ skew-symmetric}, U \in \mathbb{R}^3$$

Regression in a Riemannian manifold

But where's the regression in all of this?

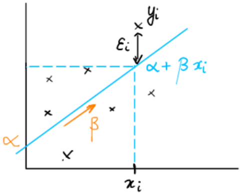
In an Euclidean space	In a Riemannian manifold
Let $X \in \mathbb{R}$ a non-random variable	Let $X \in \mathbb{R}$ a non-random variable
and $Y \in \mathbb{R}^n$ a random variable	and $Y \in \mathcal{M}$ a random variable
Regression model: $Y = \alpha + X\beta + \epsilon$	Regression model: $Y = \mathrm{Exp}\Big(\mathrm{Exp}(p, Xv), \epsilon\Big)$
α is the intercept	p is the initial point
β is the slope	v is the initial velocity vector
$(\hat{\alpha}, \hat{\beta}) = \arg\min \lvert\lvert Y - \alpha - X\beta \rvert\rvert^2$	$(\hat{p}, \hat{v}) = \arg\min d\Big(\mathrm{Exp}(p, Xv), Y\Big)^2$
A closed-form solution exists	There is usually **NO** closed-form solution

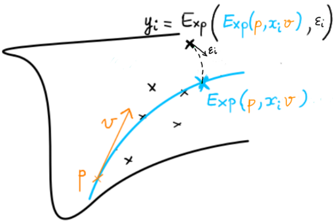




But where's the regression in all of this?

In an Euclidean space	In a Riemannian manifold
Regression model: $Y = \alpha + X\beta + \epsilon$ $(\widehat{\alpha}, \widehat{\beta}) = \arg\min \|\| Y - \alpha - X\beta \|\|^2$	Regression model: $Y = \mathrm{Exp}\Big(\mathrm{Exp}(p, Xv), \epsilon\Big)$ $(\widehat{p}, \widehat{v}) = \arg\min d\Big(\mathrm{Exp}(p, Xv), Y\Big)^2$





"Optimal" curve

- The "optimal" curve of order k is obtained by computing the parameters that minimise the least squares criterion:

$$E(\gamma) = \frac{1}{N}\sum_{j=1}^{N} d(\gamma(t_j), y_j)^2 \quad (1)$$

$$\text{under the constraint } \nabla_{\gamma'}^{(k)}\gamma'(t) = \frac{D}{dt}^{(k)}\gamma'(t) = 0 \quad (2)$$

1. Cancelling the differential form dE,
2. We take advantage of the Lie group structure by first solving in , then "transporting" the solution in G.

- developed for the Levi-Civita connection and extended to an arbitrary connection.

Simulations: Straight lines

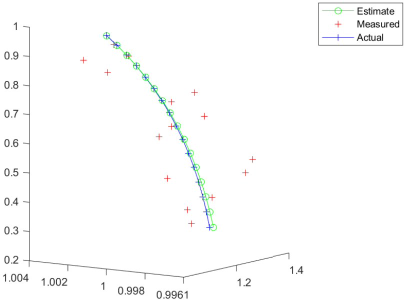
$$R^2 = 1 - \frac{SSE}{\text{Data Fréchet variance}} = 1 - \frac{\sum_j d(\gamma(t_j), y_j)^2}{\min_{y \in} \sum_j d(y, y_j)^2}$$

Maneuver	Noise std	Nmeas	R^2
Straight line	10^{-3}	25	1
Straight line	10^{-2}	**25**	**0.99**
Straight line	10^{-1}	25	0.84
Straight line	10^{-2}	25	0.99
Straight line	10^{-2}	15	0.99
Straight line	10^{-2}	10	0.99
Straight line	10^{-1}	25	0.84
Straight line	10^{-1}	15	0.82
Straight line	10^{-1}	10	0.81

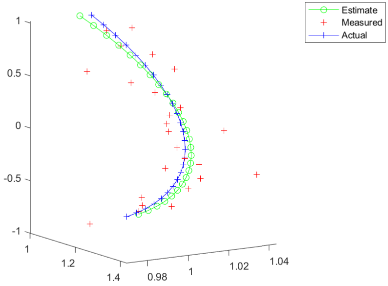
Simulations: gradual turn

Maneuver	Noise std	Nmeas	R^2
Gradual turn	10^{-2}	25	0.99
Gradual turn	10^{-2}	**15**	**0.99**
Gradual turn	10^{-2}	10	0.99
Gradual turn	10^{-1}	25	0.79
Gradual turn	10^{-1}	15	0.75
Gradual turn	10^{-1}	10	0.74
Sharp turn	10^{-1}	**25**	**0.92**
Sharp turn	10^{-1}	15	0.89
Sharp turn	10^{-1}	10	0.81

Gradual and Sharp turn

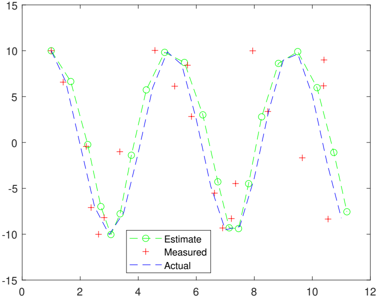


Gradual turn

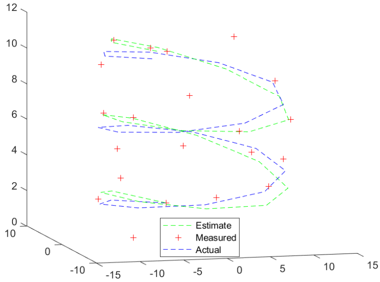


Sharp turn

Helix

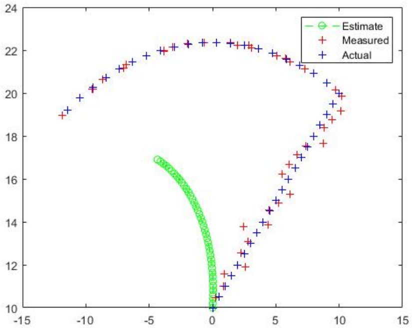


An helix in 2-D view



An helix in 3-D view

Results that do not work so much



What else?

Work in progress

- NP regression, Local polynomials
- Asymptotic properties

And more...

- Description of the noise: the Gaussian hypothesis holds only if the noise is "not too large"
- The case of unknown manifold, estimating the connection from data
- When data are more complex: time series lying in the manifold, functional data on the manifold ?

All `MATLAB` scripts and some data sets are freely available at
`https://github.com/jhnaby/LieGroupRegression`

Johan Aubray, Florence Nicol. Polynomial Regression on Lie Groups and Application to SE(3). Entropy, 2024, 26 (10), pp.825. ⟨10.3390/e26100825⟩. ⟨hal-04747804⟩