Regression on a Lie group with any connection

Johan Aubray, Florence Nicol, Stéphane Puechmorel

▶ **To cite this version:**

Johan Aubray, Florence Nicol, Stéphane Puechmorel. Regression on a Lie group with any connection. 55èmes Journées de Statistiques, Société Française de Statistique (Sfds); Université de Bordeaux, May 2024, Bordeaux (France), France. hal-04606135

HAL Id: hal-04606135
https://hal.science/hal-04606135v1

Submitted on 9 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

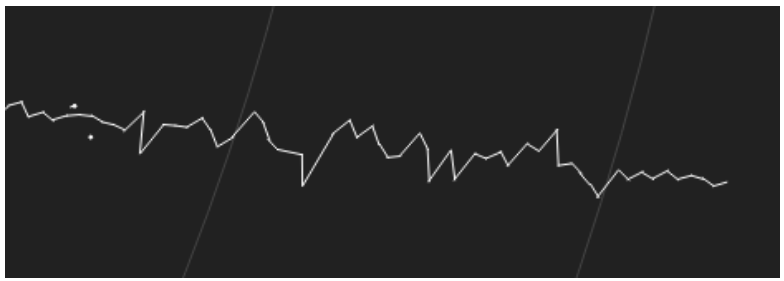
Régression sur un groupe de Lie avec une connexion quelconque

55èmes Journées de Statistique de la Sfds
Université de Bordeaux

Johan AUBRAY, Florence NICOL and Stéphane PUECHMOREL
Ecole Nationale de l'Aviation Civile, Toulouse, France
Monday May 27th, 2024

Our application

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

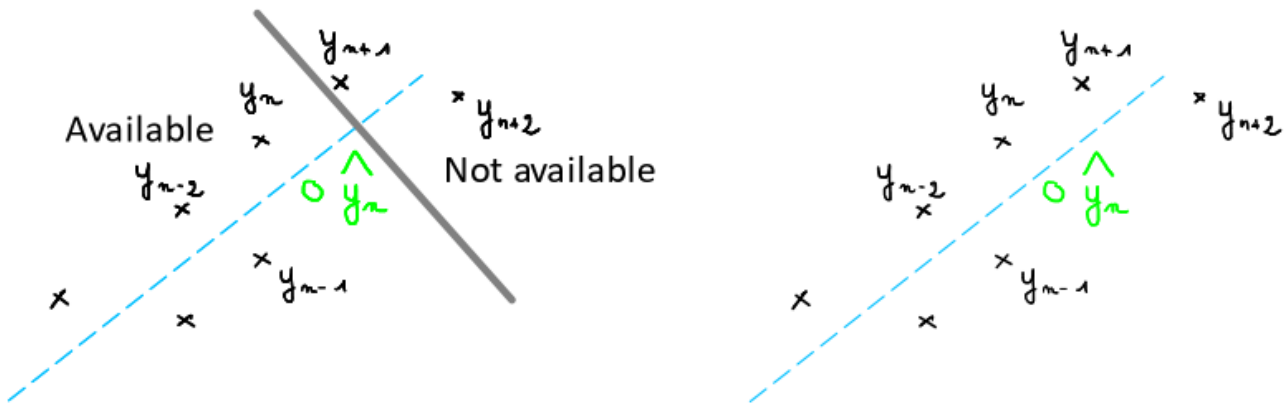


- This may be difficult with a highly maneuvering target

©2022, ENAC

Our application

- Estimating the true position of a mobile object (aircraft, drone, ...) from measured (therefore noisy) raw positions
- Two cases :
  - From past positions: prediction
    - Frequently solved with linear / extended / unscented Kalman filters, or particle filters
  - From past and future positions: smoothing

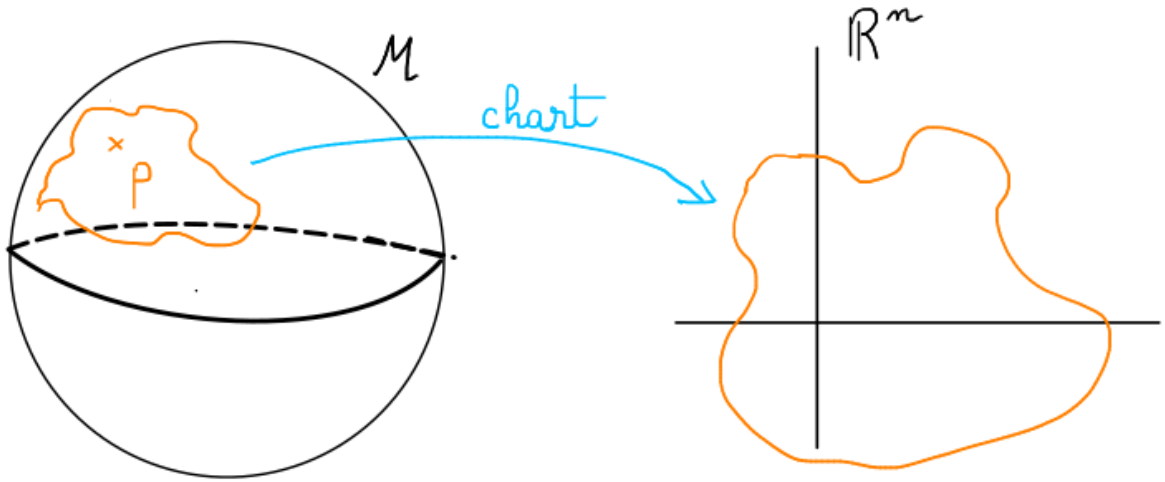


- What comes next is based on the second case: smoothing

©2022, ENAC

Our framework

- 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 $SE(3)$
- $SE(3)$ is a Lie group, which is both:
  - a differentiable manifold

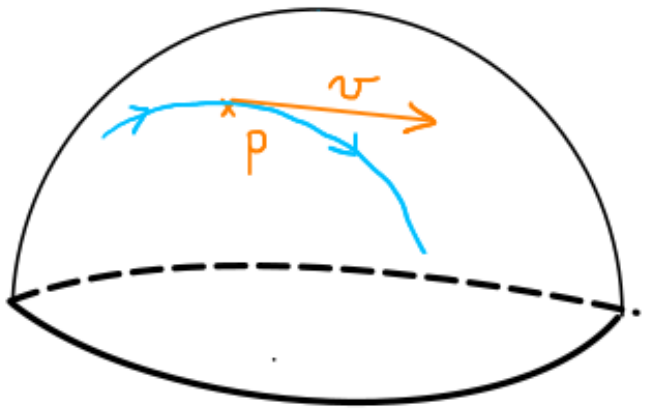


  - a group in which both the product and inverse maps are smooth

©2022, ENAC

Going off on a tangent (space)

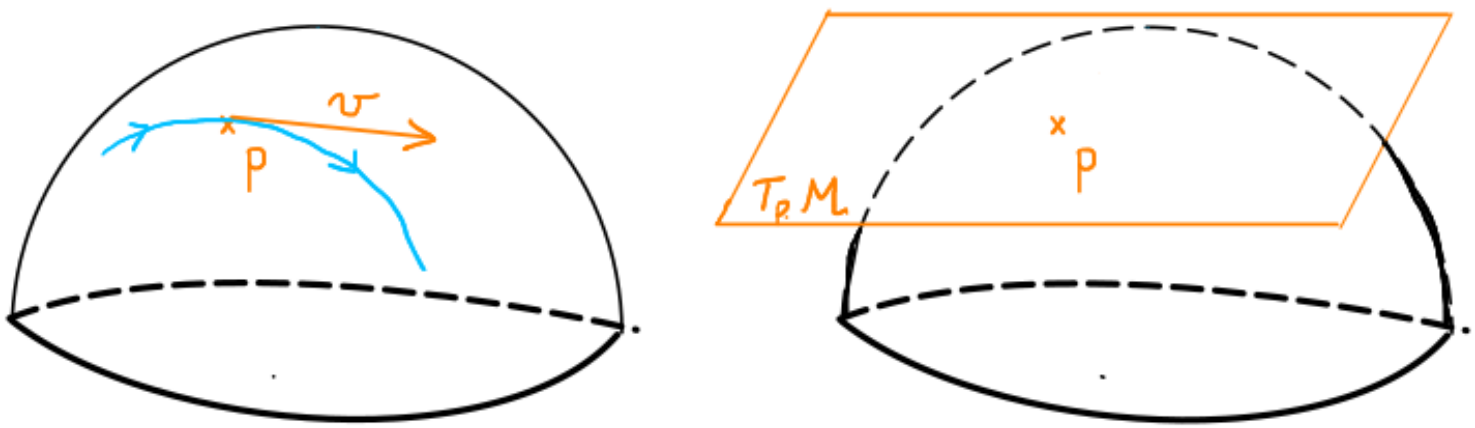
- At any point p of a manifold $\mathcal{M}$, one may define a tangent vector v as the derivative of a curve going through p: $v = c'(0)$ with $c\colon]-\epsilon, \epsilon[\to \mathcal{M}$, $c(0) = p$



©2022, ENAC

Going off on a tangent (space)

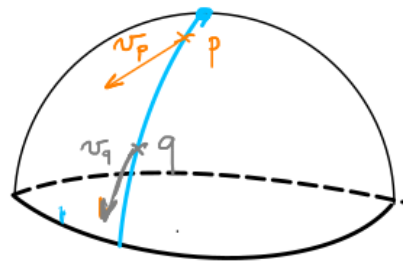
- At any point p of a manifold $\mathcal{M}$, one may define a tangent vector v as the derivative of a curve going through p: $v = c'(0)$ with $c\colon]-\epsilon, \epsilon[\to \mathcal{M}$, $c(0) = p$
- The set of all tangent vectors at p is a vector space called the tangent space at p denoted $T_p\mathcal{M}$



©2022, ENAC

Manifolds and geodesics

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



  - A connection ∇ is a rule for taking the directional derivative of a vector field along a tangent vector
  - A curve γ is called a geodesic if

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

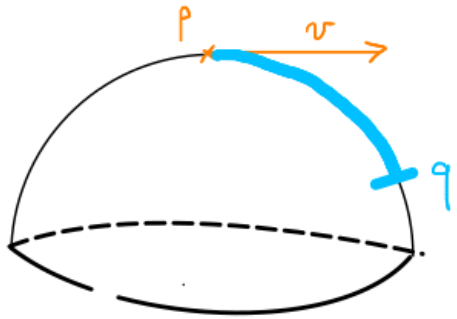
(we "parallel-transport" the $\gamma'(t)$ vector along the curve γ)

©2022, ENAC

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:

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



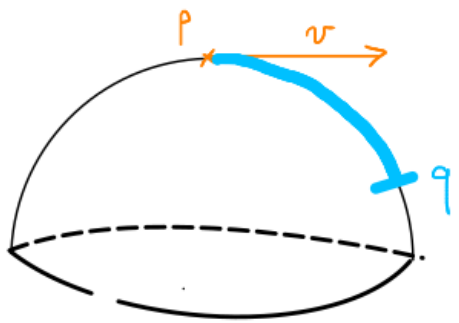
- This is the equivalent of $q = p + \overrightarrow{v}$ in $\mathbb{R}^n$

©2022, ENAC

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 $\overrightarrow{v} = \overrightarrow{pq} = q - p$ in $\mathbb{R}^n$

©2022, ENAC

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

©2022, ENAC

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

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

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

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

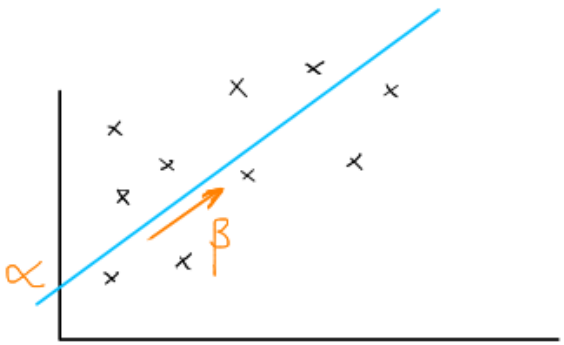
- An element of $\mathfrak{se}(3)$ can be represented by a 4-by-4 matrix:

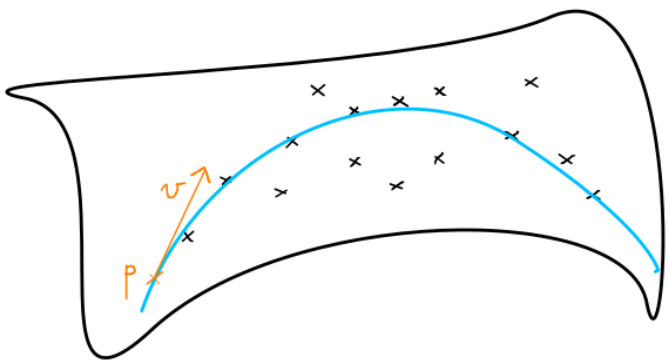
$$\begin{pmatrix} \Omega & U \\ 0 & 0 \end{pmatrix}, \Omega \text{ skew-symmetric}, U \in \mathbb{R}^3$$

©2022, ENAC

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 and $Y \in \mathbb{R}^n$ a random variable	Let $X \in \mathbb{R}$ a non-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 β is the slope	p is the initial point 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

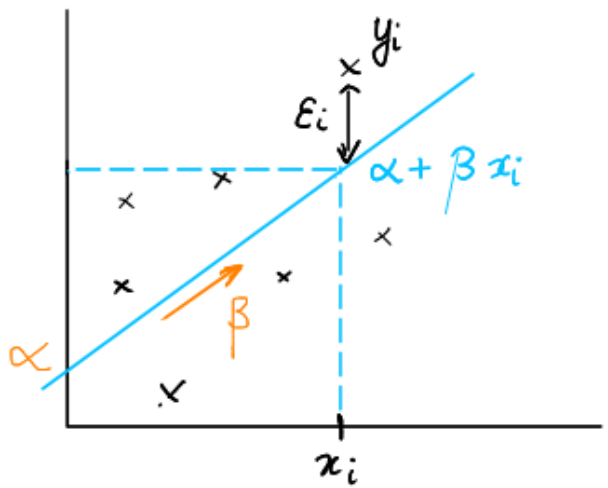


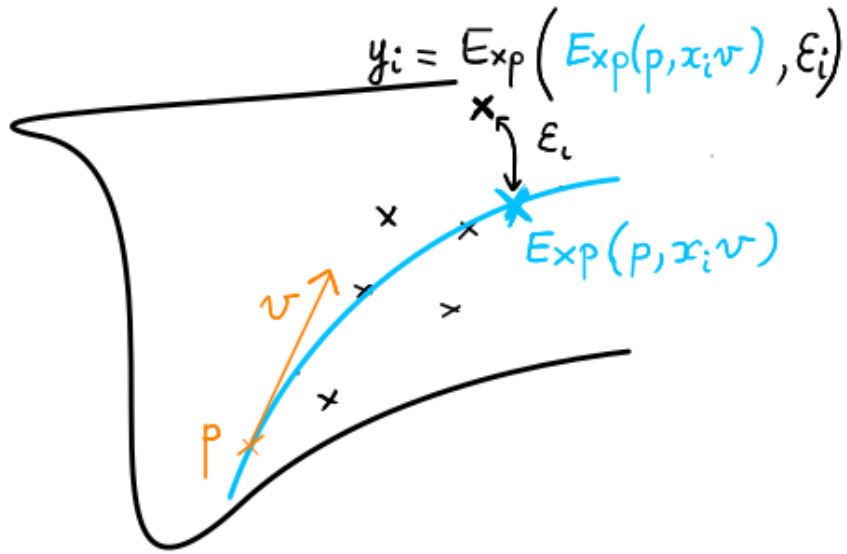


©2022, ENAC

But where's the regression in all of this?

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





©2022, ENAC

"Optimal" curve

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

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

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

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

©2022, ENAC

"Any" connection?

- There is a unique metric-compatible, torsion-free connection called the **Levi-Civita** connection ∇^{LC}
- For another connection ∇: the mapping

$$(t_1, \ldots, t_n) \in [0,1]^n \mapsto \mathrm{Exp}\left(x, \sum_{i=1}^{n} t_i e_i\right)$$

with $(e_1, \ldots, e_n)$ be a basis of $T_x\mathcal{M}$, defines a local coordinate system in a neighborhood of x

©2022, ENAC

Goodness of fit

We can use (P.T. Fletcher, 2011) a "Riemannian" coefficient of determination R^2, defined as:

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

©2022, ENAC

What's next?

- Local polynomials
- Description of the noise: the Gaussian hypothesis holds only if the noise is "not too large"

©2022, ENAC