Extrinsic Local Regression on Manifold-Valued Data

Hwiyoung Lee

Department of Statistics, Florida State University
Overview

1. Introduction

2. Extrinsic Regression

3. Applications
   3.1 VW Mean for hand shape
   3.2 VW Regression for rat data
   3.3 Grassmannian

4. Conclusion
Introduction

Let $Y \in \mathcal{M}$ be the response variable where $(\mathcal{M}, \rho)$ is a metric space, and $X \in \mathbb{R}^m$ be a predictor variable. Given data $(x_i, y_i)$

- Regression: modeling the relationship between $X$ and $Y$
  - $Y$: Manifold-valued response
  - $X$: Euclidean predictors

Typical regression framework

$$y_i = F(x_i) + \varepsilon_i$$

is not appropriate

- $y_i - F(x_i)$ is not well defined ($\mathcal{M}$ is not a vector space)
Intrinsic regression

\[ F(x) = \arg\min_{q \in \mathcal{M}} \int_{\mathcal{M}} \rho^2(q, y) P(dy|x) \]

- \( \mathcal{M} \) is a \( d \) dimensional smooth manifold
- \( \rho \) is a geodesic distance

- **Drawbacks of intrinsic model**
  - Heavy computation
    - Evaluation of an expensive gradient in an iterated algorithm
  - Sensitive to start points
    - The objective functions often have multiple modes
  - Existence and Uniqueness of the population regression function holds only under restrictive support conditions
Extrinsic Regression

\[ F(x) = \arg\min_{q \in \mathcal{M}} \int_{\mathcal{M}} ||J(q) - J(y)||^2 P(dy|x) \]

\[ = \arg\min_{q \in \mathcal{M}} \int_{\tilde{\mathcal{M}}} ||J(q) - z||^2 \tilde{P}(dz|x) \]

- **$J : \mathcal{M} \rightarrow E^D$** where $D \geq d$

- **$\tilde{\mathcal{M}} = J(\mathcal{M})$ : image of the embedding**

- **$\tilde{P}(\cdot|x) = P(\cdot|x) \circ J^{-1}$**

  : Conditional probability measure on $J(\mathcal{M})$ given $x$ induced by $P(\cdot|x)$ via the embedding $J$
Extrinsic Kernel estimate of $F(x)$

$$
\hat{F}_E(x) = J^{-1}(\mathcal{P}(\hat{F}(x))) = J^{-1}(\arg\min_{q\in\tilde{M}} ||q - \hat{F}(x)||)
$$

where,

$$
\hat{F}(x) = \arg\min_{y\in E^D} \sum_{i=1}^{n} \frac{K_H(x_i - x)||y - J(y_i)||^2}{\sum_{i=1}^{n} K_H(x_i - x)}
$$

$$
= \sum_{i=1}^{n} \frac{J(y_i) K_H(x_i - x)}{\sum_{i=1}^{n} K_H(x_i - x)}
$$
Kernel weight function

- **$K : \mathbb{R}^m \to \mathbb{R}$** is a multivariate kernel function
  - $\int_{\mathbb{R}^m} K(x) dx = 1$, $\int_{\mathbb{R}^m} xK(x) dx = 0$

- **$H = \text{Diag}(h_1, \cdots, h_m)$** with $h_i > 0$ is a bandwidth vector
  - $|H| = h_1 \cdots h_m$
  - $K_H(x) = \frac{1}{|H|} K(H^{-1}x)$
Properties

- Two step estimator
  - Step 1: Local regression on the Euclidean space after embedding
  - Step 2: Project the solution back onto the manifold

- Robust estimator using $L_1$ Euclidean norm

- Generalization using higher order local polynomial regression

Local Linear estimator

$$(\hat{\beta}_0, \hat{\beta}_1) = \arg\min_{\beta_0, \beta_1} \sum_{i=1}^{n} ||J(y_i) - \beta_0 - \beta_1^T (x_i - x)||^2 K_H(x_i - x)$$

$$\hat{F}(x) = \hat{\beta}_0(x)$$
Equivariant embedding

Since embedding $J$ is in general not unique, we need to find the optimal embedding (Equivariant Embedding).

If we can find a group homomorphism $\phi : G \rightarrow GL(D, \mathbb{R})$ s.t

$$J(gq) = \phi(g)J(q)$$

for any $g \in G$ and $q \in \mathcal{M}$

- Preserves many geometric features
- image of $\mathcal{M}$ under the group action of the Lie group $G$ is preserved by the group action of $\phi(G)$ on the image
- In some cases constructing an equivariant embedding can be a nontrivial
- In most cases a natural embedding arises and such embeddings can often be verified as equivariant
Some examples of Equivariant embeddings

1. Sphere ($S^d$) : Inclusion map

\[\begin{align*}
\iota : S^d &\rightarrow \mathbb{R}^{d+1} \\
\iota(y) &= y
\end{align*}\]

2. Planar shape $\Sigma_2^k$ : Veronese-Whitney embedding

\[\begin{align*}
j : \Sigma_2^k &\rightarrow \mathcal{S}(k, \mathbb{C}) \\
j(\sigma(z)) &= uu^*, \text{ where } ||u|| = 1
\end{align*}\]

- $u = \frac{z - <z>}{||z - <z>||}$ : preshape, $<z> = (\sum_{i=1}^{k} z_i/k, \cdots, \sum_{i=1}^{k} z_i/k)$
- $u^*$: conjugate transpose of $u$
- $j(\Sigma_2^k) = \{ A \in \mathcal{S}^+(k, \mathbb{C}) : \text{rank}(A) = 1, \text{Trace}(A) = 1, A1_k = 0\}$
Some examples of Equivariant embeddings

3 Grassmannian $G_k(\mathbb{R}^m)$: Dimitric embedding

\[ \begin{cases} 
  j : G_k(\mathbb{R}^m) \mapsto \text{Sym}(m, \mathbb{R}) \\
  j(V) = xx^\top, \text{ where } x^\top x = I_k \ (x_i \in V) 
\end{cases} \]

- map $V \in G_k(\mathbb{R}^m)$ to the matrix associated with the orthogonal projection onto $V$ with respect to $e$ where $e$ is the orthonormal basis of $\mathbb{R}^m$

- Generalizes the VW embedding ($k = 1$)
Example: Hand shape data

Hand image dataset\(^1\)

- 40 images of human hands with a resolution of 1,600 × 1,200
- 4 people contributed with 10 images each of their left hand
- 56 landmarks

\(^1\)http://www.imm.dtu.dk/~aam/
Registration via VW embedding

**Figure:** landmark image of unregistered data

**Figure:** landmark image of registered data via VW embedding
For $\Sigma^k_2$, the Procrustes mean agree with the VW mean [Bhattacharya and Patrangenaru, 2003]
Example $\Sigma^k_2$: Rat Calivarium Growth data

- First analyzed by Bookstein [Bookstein, 1991]
- $k = 8$ landmarks on a midsagittal section of rat calvaria
- Landmark positions are available for 18 rats and at 8 ages apiece

Figure: landmark images: unregistered data (left), Preshape (right)
Extrinsic mean via VW embedding

Figure: VW means of age groups
Extrinsic regression via VW embedding

Figure: VW regression of rat calvarium ($h = 1$)
VW regression

Figure: VW regressions with different bandwidth (h = 1(black dashed), 10(red solid), 100(blue dotted)) : 4th(left), 5th(middle), and 7th landmark(right)

- small $h$ captures local behavior
Example $G_k(\mathbb{R}^m)$: Facial Attractiveness data

- 86 × 86 pixel images
- taken from http://www.hotornot.com
- Rectified with affine transformation so that common landmarks (eyes, nose, corners of the mouth) are in canonical locations
- Data gathered by White et al.
  http://www.ryanmwhite.com/research/tr_hot.html
Eigenface [Sirovich and Kirby, 1987]

The original image can be represented as the linear combination of the eigen face

\[-1.9726 \times 1 \text{st eigen face} + 0.3016 \times 2 \text{nd eigen face} + 0.3644 \times 3 \text{rd eigen face} + 0.0580 \times 4 \text{th eigen face} \cdots\]

Result

- first 50
- first 100
- first 200
- first 1000
Local Polynomial regression

Local polynomial regression result of female data: top $(h = 0.05)$, bottom $(h = 1)$ evaluation points (attractive score: 4, 5, 6, 9 from the left)
Conclusion

Intrinsic vs Extrinsic

- We dealt with manifold for which equivariant embeddings are available
- Extrinsic approaches are in general advantageous over the intrinsic models
- There are complex manifolds such as higher dimensional shape spaces for which good embedding is hard to construct
- We expect intrinsic models to perform better than extrinsic ones
Future work

- Regression problem which has manifold valued covariate

\[ y_i = F(x_i) + \varepsilon_i \]

- \(X \in \mathcal{M}\)

- \(Y \in \mathbb{R}\)
  Attractiveness, age, \ldots
Reference

