Nonlinear Regression on Riemannian Manifolds and Its Applications to Neuro-Image Analysis⋆

Monami Banerjee[1], Rudrasis Chakraborty[1], Edward Ofori[2], David Vaillancourt[2], and Baba C. Vemuri[1,⋆⋆]

[1] Department of CISE, University of Florida, Gainesville, Florida, USA
{monami,rudrasis,vemuri}@cise.ufl.edu
[2] Department of Applied Physiology and Kinesiology, University of Florida, Florida, USA
{eofori,vcourt}@ufl.edu

Abstract. Regression in its most common form where independent and dependent variables are in $\mathbb{R}^n$ is a ubiquitous tool in Sciences and Engineering. Recent advances in Medical Imaging has lead to a wide spread availability of manifold-valued data leading to problems where the independent variables are manifold-valued and dependent are real-valued or vice-versa. The most common method of regression on a manifold is the geodesic regression, which is the counterpart of linear regression in Euclidean space. Often, the relation between the variables is highly complex, and existing most commonly used geodesic regression can prove to be inaccurate. Thus, it is necessary to resort to a non-linear model for regression. In this work we present a novel Kernel based non-linear regression method when the mapping to be estimated is either from $M \rightarrow \mathbb{R}^n$ or $\mathbb{R}^n \rightarrow M$, where M is a Riemannian manifold. A key advantage of this approach is that there is no requirement for the manifold-valued data to necessarily inherit an ordering from the data in $\mathbb{R}^n$. We present several synthetic and real data experiments along with comparisons to the state-of-the-art geodesic regression method in literature and thus validating the effectiveness of the proposed algorithm.

1 Introduction

Regression is an essential tool for quantitative analysis to find the relation between independent and dependent variables. Here, we are given a training set of both of these variables and we seek a relation between them. When, both of these variables are in Euclidean space, and there is a linear relation between them, i.e., $y_i = ax_i + b$ for a set of $\{x_i, y_i\}$, a common way to solve for the unknowns a and b is using linear least-square estimator, i.e., minimizing the sum of square distances between the two sets of variables over the training set. But, in many real applications, the relation is seldom linear, hence a non-linear least squares estimator or any other sophisticated regression tool like Support Vector Regression [4] can be used.

Often, either of the independent or dependent variables are manifold-valued and lie on a smooth Riemannian manifold. In such instances, *embedding* the manifold valued

---

variables in Euclidean space (using the Whitney Embedding [1]) might result in a poor estimation of the underlying model. Also, as any general manifold globally lacks the vector space structure, any linear combination of points on the manifold may not lie on the manifold. For example, suppose the data points lie in a Kendall's shape space [14], then an arbitrary linear combination of the shapes will not yield a point on in the shape space. These problems motivate the development of novel regression methods for manifold-valued data. We will now briefly present earlier work that addresses this problem.

Related Work: Curve fitting on Riemannian manifolds where some notion of ordering is imposed on the manifold-valued data has been quite common lately in literature [2,18,7,13,5,16]. We will present a brief review within the limited space. Samir et al. [18] developed a gradient descent algorithm for time ordered manifold-valued data using a variational formulation, where the cost function entails a data fidelity and a regularization constraint on the curve being sought. This formulation itself is quite common to finding smooth approximation of both real-valued and manifold-valued data. What is then different between methods is the kind of metric used and at times even the data fidelity terms. Each could facilitate the solution sought from an efficiency and/or accuracy.

In the recent past, several researchers [7,13] have proposed geodesic regression on manifolds, as well as non-parametric regression models [2]. The geodesic regression models correspond to linear regression in $\mathbb{R}^n$. Most recently however, a variational spline regression for the manifold of diffeomorphisms was presented in a large deformation diffeomorphic mapping (LDDMM) setting [19]. Fletcher [7] proposed *geodesic regression* to regress manifold-valued data against the real-valued variables. Taking cues from [7], authors in [5], developed a regression technique for points that lie on unit Hilbert sphere. In [2], authors estimate the correlation between shape and age using manifold regression. The aforementioned methods dealt

Fig. 1. Examples of nonlinear & geodesic regression.

mostly with the independent scalar variable. A multivariate general linear model was proposed in [16] where given a dataset, authors try to model a functional relation from a $\mathbb{R}^n$ to a manifold $\mathcal{M}$. In [15], they extend Canonical Correlation Analysis (CCA) on Riemannian manifold, where both of the variables are manifold-valued. Hong et al. [12] proposed a shooting spline formulation to regress points on Grassmann manifold with reals. In [9], Hinkle et al. has proposed a polynomial regression method formulated as a variational minimization problem on the manifold using covariant derivatives. The minimization tends to covariant differential equations.

In this paper, we present a nonlinear kernel regression technique to handle both of the following commonly encountered cases, $\mathbb{R}^n \rightarrow \mathcal{M}$ and $\mathcal{M} \rightarrow \mathbb{R}^n$. We dub our proposed kernel based regression from $\mathbb{R}^n \rightarrow \mathcal{M}$ as *Manifold-valued Kernel Regression* (MVKR). A key advantage of this approach is that there is no requirement for the manifold-valued data to necessarily inherit an ordering from the multi-variate data in $\mathbb{R}^n$, a necessary requirement in most existing methods. An example in Fig. 1 depicts the usefullness in terms of accuracy in using the nonlinear regression over the geodesic regression model.
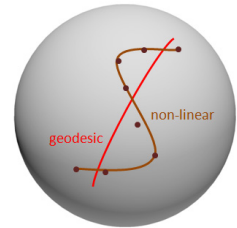
2   Methodology

Regression is ubiquitous in scientific analysis where given a set of tuples $\{x_i, y_i\}_{i=1}^N \subset \mathcal{X} \times \mathcal{Y}$, the goal is to find a functional relation between $\{x_i\}_{i=1}^N$ and $\{y_i\}_{i=1}^N$. Here, one variable is the observed data (*independent variable*) and the other one is the response (*dependent variable*). We propose a kernel interpolation to find the relation between observed data and responses where one of them lies in the Euclidean space and the other one lies on a Riemannian manifold. Given $\{x_i\}_{i=1}^N \subset \mathbb{R}^n$ and $\{y_i\}_{i=1}^N \subset \mathcal{M}$, we pose the two cases as following interpolation problems:

- **Manifold valued independent variable:** Find a function $f : \mathcal{M} \to \mathbb{R}^n$ such that $x_i = f(y_i), \forall i$.
- **Manifold valued dependent variable:** Find a function $h : \mathbb{R}^n \to \mathcal{M}$ such that $y_i = h(x_i), \forall i$.

In both the above cases, $\mathcal{M}$ is a Riemannian manifold equipped with a Riemannian metric g. We will address these above two problems separately in the following subsections.

2.1   Manifold Valued Independent Variable

Given $\{x_i, y_i\}_{i=1}^N$ as before, we try to model the function $\hat{f} : \mathcal{M} \to \mathbb{R}^n$ by minimizing the following error function: $E = \frac{1}{N} \sum_{i=1}^N \|\hat{x}_i - x_i\|^2$ where, $\hat{x}_i = \hat{f}(y_i) = \sum_{j=1}^k \mathcal{K}(c_j, y_i) t_j$. Here $\{c_j\}_{j=1}^k \subset \mathcal{M}$ and $\{t_j\}_{j=1}^k \subset \mathbb{R}^n$ are the *representatives* on $\mathcal{M}$ and $\mathbb{R}^n$ respectively. $\mathcal{K} : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$ is the *kernel function*. Thus, $\hat{x}$, the approximation of x is the weighted mean of t_j's. The weights here are computed by using a suitable kernel function and representatives, $\{c_j\}_{j=1}^k$, on the manifold, $\mathcal{M}$. We learn the $\{t_j\}_{j=1}^k$ by minimizing the above error function, E, whereas, $\{c_j\}_{j=1}^k$ are taken to be the cluster representatives. Here, we used the steepest descent technique to estimate $\{t_j\}_{j=1}^k$. The gradient of the objective function with respect to t_j is given by, $\nabla_{t_j} E = \frac{2}{N} \sum_{i=1}^N (\hat{x}_i - x_i) \mathcal{K}(c_j, y_i)$.

Note that, as the objective function, E is convex in t_j, the global minimum can be achieved using a steepest descent technique. In a similar fashion, we can initialize c_j to be the cluster representatives and estimate them along the gradient direction. The gradient of the objective function with respect to c_j is given by, $\nabla_{c_j} E = \frac{2\,t_j}{N} \sum_{i=1}^N (\hat{x}_i - x_i) \nabla_{c_j} \mathcal{K}(c_j, y_i)$.

Since any kernel function depends on the underlying metric, if the underlying manifold $\mathcal{M}$ has a closed form expression for the geodesic distance, so will $\nabla_{c_j} \mathcal{K}(c_j, y_i)$. In this work, we use the kernel $\mathcal{K}(c, y) = \exp\{-\frac{b}{2\sigma^2} d(c, y)^2\}$, where b, σ^2 are the kernel parameters, and $d(., .)$ is the geodesic distance on $\mathcal{M}$. Then, $\nabla_{c_j} \mathcal{K}(c_j, y_i) = \frac{b}{2\sigma^2} \mathcal{K}(c_j, y_i) Log_{c_j} y_i$ where, $Log_{c_j} y_i$ is the Riemannian inverse exponential map. Note that, the b value is tuned according to the structure of the dataset. By drawing an analogy with the Gaussian kernel on $\mathbb{R}^n$, we chose a small b value for a well clustered data, and a high b value otherwise. The parameter σ^2 is taken as the variance over the training data.

2.2   Manifold Valued Dependent Variable

Given $\{x_i, y_i\}_{i=1}^N$ as above, we now try to model the function $\hat{h} : \mathbb{R}^n \to \mathcal{M}$ such that $y_i \approx \hat{h}(x_i)$. As before, let $\mathcal{M}$ be equipped with a Riemannian metric g. Also, let $d : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$ be the geodesic distance on $\mathcal{M}$ defined as follows: $d(y_i, y_j)^2 = g_{y_i}(Log_{y_i} y_j, Log_{y_i} y_j)$, where $Log_{y_i} y_j$ is the inverse-exponential map. We can now estimate h by minimizing the following error function: $E = \frac{1}{N} \sum_{i=1}^N d(\hat{y}_i, y_i)^2$ where,

$$\hat{y}_i = \hat{h}(x_i) = \arg\min_{\mu \in \mathcal{M}} \sum_{j=1}^k \mathcal{K}_{Euc}(t_j, x_i)\, d(c_j, \mu)^2 \qquad (1)$$

Analogous to the manifold valued independent variable case, here $c_j \in \mathcal{M}$ and $t_j \in \mathbb{R}^n$, $\forall j$. $\mathcal{K}_{Euc} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is the kernel function on the Euclidean space. Thus, y_i is estimated as the weighted Fréchet mean (FM) [8] of the representatives, $\{c_j\}_{j=1}^k$, where weights are given by the kernel function, yielding the MVKR. We use $\{t_j\}_{j=1}^k$ as the cluster representatives and estimate $\{c_j\}_{j=1}^k$ using the steepest descent on the objective function. The gradient direction of E with respect to c_j is given by,

$$\nabla_{c_j} E = -\frac{2}{N} \sum_{i=1}^N Log_{\hat{y}_i} y_i \, \nabla_{c_j} \hat{y}_i. \qquad (2)$$

As c_j and $\hat{y}_i$ both are on $\mathcal{M}$, we will use charts to compute $\nabla_{c_j} \hat{y}_i$. Let $\mathcal{M}$ be an m dimensional manifold. Consider two charts (U, Φ) and (V, Ψ) containing c_j and $\hat{y}_i$, respectively. By fixing x_i, we can take $\hat{y}_i$ as a function of c_j's. Let the function be F. Then, $\nabla_{c_j} \hat{y}_i$ can be defined as $\nabla_{c_j} \hat{y}_i := \nabla_{\tilde{c}_j} G$, where $\tilde{c}_j = \Phi(c_j)$ and $G = \Psi \circ F \circ \Phi^{-1} : \mathbb{R}^m \to \mathbb{R}^m$. Hence, $\nabla_{\tilde{c}_j} G$ is the Jacobian of G. Note that, $\nabla_{c_j} E \in T_{c_j} \mathcal{M}$, so in order to make the RHS of equation 2 to be in $T_{c_j} \mathcal{M}$, we use parallel transport of $Log_{\hat{y}_i} y_i$ from $\hat{y}_i$ to c_j. For a general Riemannian manifold $\mathcal{M}$, we can approximate this parallel transport, $\Lambda_{c_j} Log_{\hat{y}_i} y_i$ as $\Lambda_{c_j} Log_{\hat{y}_i} y_i \approx Log_{c_j} y_i - Log_{c_j} \hat{y}_i$.

Since there is no closed form solution for the weighted FM of more than two samples on general Riemannian manifolds, computation of $\nabla_{\tilde{c}_j} G$, or the Jacobian of G, is not straightforward. Hence, in spirit of [16,11], we approximate Equation 1 as $\hat{y}_i \approx Exp_p \left(\sum_{j=1}^k \mathcal{K}_{Euc}(t_j, x_i) Log_p c_j \right)$, where $p \in \mathcal{M}$ is any arbitrary point on M, and Exp is the Riemannian Exponential map. In the absence of such an approximation, the problem would become analytically intractable as estimating both the control points and the FM jointly is nontrivial. With this simplification, $\nabla_{c_j} \hat{y}_i = \mathcal{K}_{Euc}(t_j, x_i) \times I_m$, where I_m is the identity matrix of size m. For the case of $P(n)$, we resort to use of the efficient recursive FM estimator in [10] and a similar one for and S^n.

3   Experimental Results

We now evaluate the performance of the proposed regression method on both synthetic and real datasets. In the following two subsections, we will experimentally show effectiveness of our method to (1) regress real vector-valued dependent variables against

manifold-valued independent variables and (2) regress manifold-valued dependent variables against real vector-valued independent variables. In order to quantify the performance of our $\mathbb{R}^n$ to manifold regression, we use the R^2 statistical measure and the $p-$value. The R^2 statistical measure on a manifold is defined in [7] and repeated here for convenience. Let $\{y_i\}_{i=1}^N$ be the manifold-valued data with its corresponding predicted value to be $\{\hat{y}_i\}_{i=1}^N$. Let the *unexplained variance* be defined as $\sum_{i=1}^N d(y_i, \hat{y}_i)^2$. Then, the R^2 statistic is defined as: $R^2 = 1 - \frac{\text{unexplained variance}}{\text{data variance}}$. The value of R^2 statistic lies in the interval $[0, 1]$, and a value close to one in general denotes better regression performance. We use a $t-$test over 30 independent runs to reject the null hypothesis, H_0: *mean of the unexplained variance is not less than the mean of the data variance* with a significance level of 0.001. For the manifold to $\mathbb{R}^n$ regression, we present an application to the classification on Parkinson's dataset and report the average classification accuracy over 30 runs.

3.1   Manifold Valued Independent Variable

In this section, we present results of our regression scheme applied to classification of MR T2 brain scans obtained from, (1) controls (CON), and patients with (2) essential tremor (ET), and (3) Parkinson's disease (PD). We aim to automatically discriminate between these three classes, using features derived from the data.

In [20], authors have used DTI based analysis, specifically the scalar-valued features to address the problem of movement disorder classification. In this section, we use the shape of the Substatia Nigra across the input population as our key discriminatory feature. Sample Substantia Nigra shapes for the three classes

Fig. 2. Examples of Substantia Nigra

are shown in Fig. 2. The shapes of interest are first segmented and then are converted into a probability density function. Then using the square root density parameterization, this shape can be represented as a point on the unit Hilbert sphere using the Schrodinger Distance Transform (SDT) [3].
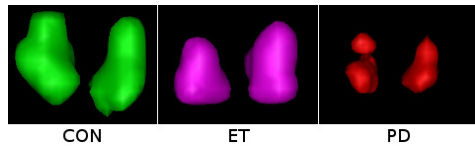
The key feature used in our classification of the aforementioned disease classes is the shape of the Substantia Nigra. The Substantia Nigra was hand-segmented from all rigidly pre-aligned datasets, consisting of 25 controls, 15 ET and 24 PD images. The T2 brain scans were acquired using a 3T Phillips MR scanner with the following parameters: $TR = 774\,ms$, $TE = 86\,ms$ and voxel size $= 2 \times 2 \times 2\,mm^3$.

We first collected random (point) samples on the boundary of each 3-D Substantia Nigra shape, and applied the SDT to represent each shape as a point on the unit hypersphere. The size of the ROI for the 3-D shape of interest was set to $(28 \times 28 \times 15)mm^3$, resulting in a 11760-dimensional unit vectors using SDT. Therefore, the samples now live on the $\mathbb{S}^{11759}$ manifold.

We randomly selected 10 Control, 10 PD and 5 ET images as the test set, and used the rest of the data for training. The details of our classification method are described next. First, we regress the dependent variable against the independent variable on $\mathbb{S}^{11759}$. In

order to make the dependent variable lie in $[0, 1]$, we apply the logistic function $\mathcal{L}$ on the dependent variable $f(y)$. Then, we classify a point y as belonging to class-1, if $\mathcal{L}(f(y)) < 0.5$, else we assign it to class-2. The classification task is repeated 30 times using various randomly chosen training sets and the average accuracy is reported. The results are shown in Table 1. We compare our method with the standard PCA and PGA (Principal Geodesic Analysis) [6], and report the accuracy of classification.

The results show that our proposed method performs well compared to the other two in classifying Control versus PD and ET. In case of PD vs. ET

Table 1. Result based on Substantia Nigra shape

	Control vs. ET			Control vs. PD			PD vs. ET		
	Proposed	PGA	PCA	Proposed	PGA	PCA	Proposed	PGA	PCA
Accuracy	**100.00**	90.14	75.69	**95.26**	92.95	67.32	85.71	**87.58**	64.60

classification, our method gives slightly lower accuracy compared to PGA.

3.2 Manifold Valued Dependent Variable

In this section, we applied our MVKR method on synthetic and real datasets. In all of these experiments, we have made a comparison with the recently proposed MGLM method in [16] and MKRE (Manifold kernel regression estimator) in [2]. As MVKR and MKRE both use the same Nadaraya-Watson kernel, we have used the same choice of parameters for both of these methods.

Synthetic Data Experiment: For this experiment, we synthesized a dataset $\{\mathbf{x}_i, y_i\}_{i=1}^{500} \subset \mathbb{R}^2 \times S^2$ by defining a function $h : \mathbf{x}_i = [\theta_i, \phi_i] \to y_i$ as follows: $h([\theta_i, \phi_i]) := (\cos(\theta_i) \cos(\phi_i),$ $\cos(\theta_i) \sin(\phi_i), \sin(\theta_i))$ where $\theta_i \in [0, \pi/2)$, $\phi_i \in [0, 2\pi]$, $\forall i$. Thus, all the y_is are on the northern

Table 2. Synthetic data results

	MVKR	MGLM	MKRE
Train Error	**0.00**	0.60	0.07
Test Error	**0.00**	0.61	0.07
R^2 Stat.	**1.00**	0.29	0.92
$p-$value	**< 0.001**	< 0.001	< 0.001

hemisphere of the $2-$sphere, so FM is uniquely defined. We have partitioned this dataset into 90%, 10% for training and testing respectively. The $p-$value and average R^2 statistics are reported in Table 2 over 30 runs. From these figures, we can clearly see that our MVKR method performs better in comparison to MGLM [16] and gives comprative performance to MKRE [2].

OASIS Dataset [17]: We used the publicly available OASIS data [17] to regress manifold-valued data with reals. This dataset consists of T1 MR brain scans of subjects with ages from 18 to 96 including individuals with early stage Alzheimer's Disease.

We randomly chose 4 brain scans from each of the decades in the $18 - 96$ age group, totalling 36 brain images, out of which 32 were randomly chosen and used as training and the rest were used as the test set. Corpus callosum (CC) shapes of individuals of varying ages are shown in Fig. 3. We seek to model the relationship between age and shape of the CC, captured using three

Fig. 3. Corpus callosum shapes

different features as described in the following. From each of the brain images of the 36 individuals, we construct three different data representations as follows. (1) We segmented out the CC from the brain images. Then, we take the boundary of the CC and
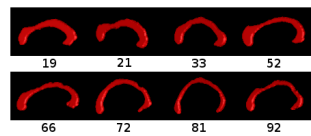
map it to S^{24575} using the SDT [3]. (2) After segmenting out the CC, we used a set of landmark points on the boundary and map each of these point sets into the Kendall's shape space [14], which is a *complex projective space*. (3) We took the whole brain image and computed the normalized histogram and used the square root of the normalized histogram to map each image on to S^{255}.

The average R^2 statistics of 30 runs on each of these three representations of the chosen OASIS datasets is given in Table 3. From the table, it is evident that the performance of MVKR is significantly better compared to the MGLM method. It should be noted that the R^2 statistics reported by MVKR is not very high (not close to 1). But it can be argued that, as we are only considering relation between age and the manifold-valued data, the relation is highly nonlinear.

Hence, it is not possible to truly capture the "relation in full" based on age alone, of an individual. Also, the brain images are chosen randomly

Table 3. Results on the OASIS dataset

	Dataset using SDT			Kendall's shape space			Dataset using histogram		
	MVKR	MGLM	MKRE	MVKR	MGLM	MKRE	MVKR	MGLM	MKRE
R^2 Stat.	**0.49**	0.05	0.46	**0.35**	-0.27	0.33	**0.48**	-0.18	0.40
$p-$value	**< 0.001**	< 0.001	< 0.001	**< 0.001**	> 0.001	< 0.001	**< 0.001**	> 0.001	< 0.001

without considering gender, educational background or even symptoms of AD, all of which makes the relation between age and the shape of the CC very complex. So, given these confounding parameters that could influence the structure, the R^2 statistics for MVKR depicts a significantly good performance. Note that, for second and third variant of this dataset, MGLM results in a negative R^2 statistic. From the definition of R^2 statistics, we can see that a negative value indicates that the regressor performed worse than the most trivial choice, which is FM of the dataset for any given test point x (value of the independent variable).

Thus, MGLM's unsatisfactory performance on these datasets indicates that a linear regressor is inept for this problem and motivates the use of a nonlinear regression technique such as the one presented here. The $p-$values reported in Table 3 indicate the higher statistical significance and hence the superior performance of our MVKR method. In comparison to MKRE, the performance of MVKR is consistently better, though not by a significant amount.

So, in summary, as for most of the real cases, the data on the manifold do not lie close to a geodesic, the performance of MGLM is not comparable to MVKR. This is due to the fact that MGLM assumes that data lie close to a geodesic while MVKR does not require any such assumption. When the data lie or are close to a geodesic, MVKR and MGLM have comparable performance as can be seen from the following toy example. In this example, we have used the sythetic data on $P(3)$, the space of symmetric positive definite matrices, in [16]. The R^2 statistics value for MGLM and MVKR are 0.98 and 0.99 respectively. We would also like to point out that although compared to MKRE, performance of MVKR is not significantly better, MVKR is applicable for $\mathbb{R}^n$ to M regression and vice versa, whereas, the method in [2] is applicable only to regression for the case of $\mathbb{R}$ to M.

4   Conclusions

In this paper, we presented a novel nonlinear regression technique for estimating the functional relationship between manifold-valued independent variables and $\mathbb{R}^n$ valued dependent variables and vice versa. Earlier work in this area involved use of geodesic regression and is ill suited for many situations involving complex relationships between the aforementioned independent and dependent variables. Our method involved a Kernel-based technique and we presented several experiments to demonstrate the performance of our methods in comparison to the state-of-the-art (MGLM method) on a variety of data sets. Results depict that our method yields superior performance for both the applications namely, classification of movement disorders and finding a correlation between age and CC shape of patients from the OASIS database.

References

1. Adachi, M., Hudson, K.: Embeddings and immersions. American Mathematical Soc. (2012)
2. Davis, B.C., Fletcher, P.T., et al.: Population shape regression from random design data. In: IEEE ICCV, pp. 1–7 (2007)
3. Deng, Y., Rangarajan, A., et al.: A Riemannian framework for matching point clouds represented by the Schrodinger distance transform. In: IEEE CVPR, pp. 3756–3761 (2014)
4. Drucker, H., Burges, C.J., et al.: Support vector regression machines. In: NIPS, pp. 155–161 (1997)
5. Du, J., Goh, A., et al.: Geodesic regression on orientation distribution functions with its application to an aging study. Neuroimage, 416–426 (2014)
6. Fletcher, P., Lu, C., et al.: Principal geodesic analysis for the study of nonlinear statistics of shape. In: IEEE TMI, pp. 995–1005 (2004)
7. Fletcher, P.T.: Geodesic regression and the theory of least squares on Riemannian manifolds. International Journal of Computer Vision, 171–185 (2013)
8. Fréchet, M.: Les éléments aléatoires de nature quelconque dans un espace distancié. Annales de l'institut Henri Poincaré, pp. 215–310 (1948)
9. Hinkle, J., Muralidharan, P., Fletcher, P.T., Joshi, S.: Polynomial regression on Riemannian manifolds. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part III. LNCS, vol. 7574, pp. 1–14. Springer, Heidelberg (2012)
10. Ho, J., Cheng, G., et al.: Recursive Karcher expectation estimators and geometric law of large numbers. In: AISTATS, pp. 325–332 (2013)
11. Ho, J., Xie, Y., et al.: On a nonlinear generalization of sparse coding and dictionary learning. In: ICML, pp. 1480–1488 (2013)
12. Hong, Y., Kwitt, R., Singh, N., Davis, B., Vasconcelos, N., Niethammer, M.: Geodesic regression on the Grassmannian. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014, Part II. LNCS, vol. 8690, pp. 632–646. Springer, Heidelberg (2014)
13. Hong, Y., Singh, N., Kwitt, R., Niethammer, M.: Time-warped geodesic regression. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) MICCAI 2014, Part II. LNCS, vol. 8674, pp. 105–112. Springer, Heidelberg (2014)
14. Kendall, D.: A survey of the statistical theory of shape. Stat. Science, 87–99 (1989)
15. Kim, H.J., Adluru, N., Bendlin, B.B., Johnson, S.C., Vemuri, B.C., Singh, V.: Canonical Correlation analysis on Riemannian Manifolds and its Applications. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014, Part II. LNCS, vol. 8690, pp. 251–267. Springer, Heidelberg (2014)

16. Kim, H.J., Bendlin, B.B., et al.: MGLM on Riemannian manifolds with applications to statistical analysis of diffusion weighted images. In: IEEE CVPR, pp. 2705–2712 (2014)
17. Marcus, D.S., Wang, T.H., et al.: OASIS: cross-sectional mri data in young, middle aged, nondemented, and demented older adults. Journal of Cognitive Neuroscience, 1498–1507 (2007)
18. Samir, C., Absil, P.-A., et al.: A gradient-descent method for curve fitting on Riemannian manifolds. Foundations of Comput. Mathematics, 49–73 (2012)
19. Singh, N., Niethammer, M.: Splines for diffeomorphic image regression. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) MICCAI 2014, Part II. LNCS, vol. 8674, pp. 121–129. Springer, Heidelberg (2014)
20. Vaillancourt, D., Spraker, M., et al.: High-resolution diffusion tensor imaging in the substantia nigra of de novo parkinson disease. Neurology, 1378–1384 (2009)