Polynomial Regression on Riemannian Manifolds

Jacob Hinkle, Prasanna Muralidharan, P. Thomas Fletcher, and Sarang Joshi

SCI Institute, University of Utah
72 Central Campus Drive, Salt Lake City, UT 84112
{jacob,prasanna,fletcher,sjoshi}@sci.utah.edu

Abstract. In this paper we develop the theory of parametric polynomial regression in Riemannian manifolds. The theory enables parametric analysis in a wide range of applications, including rigid and non-rigid kinematics as well as shape change of organs due to growth and aging. We show application of Riemannian polynomial regression to shape analysis in Kendall shape space. Results are presented, showing the power of polynomial regression on the classic rat skull growth data of Bookstein and the analysis of the shape changes associated with aging of the corpus callosum from the OASIS Alzheimer's study.

1 Introduction

The study of the relationship between measured data and known descriptive variables is known as the field of regression analysis. As with most statistical techniques, regression analyses can be broadly divided into two classes: parametric and non-parametric. The most widely known parametric regression methods are linear and polynomial regression in Euclidean space, wherein a linear or polynomial function is fit in a least-squares fashion to observed data. Such methods are the staple of modern data analysis. The most common non-parametric regression approaches are kernel-based methods and spline smoothing approaches which provide much more flexibility in the class of regression functions. However, their non-parametric nature presents a challenge to inference problems; if, for example, one wishes to perform a hypothesis test to determine whether the trend for one group of data is significantly different from that of another group.

Fundamental to the analysis of anatomical imaging data within the framework of computational anatomy is the analysis of transformations and shape which are best represented as elements of Riemannian manifolds. Classical regression suffers from the limitation that the data must lie in a Euclidean vector space. When data are known to lie in a Riemannian manifold, one approach is to map the manifold into a Euclidean space and perform conventional Euclidean analysis. Such extrinsic or coordinate-based methods inevitably lead to complications [1]. For example, given simple angular data, averaging in the Euclidean setting leads to nonsensical results which depend on the chosen coordinates; in the usual coordinates, the mean of 359 degrees and 1 degree is 180 degrees. In the case of shape analysis, regression against landmark data in a Euclidean setting, even after Procrustes alignment (an intrinsic manifold-based technique), will introduce scale and rotation in the regression function, and hence does not capture

A. Fitzgibbon et al. (Eds.): ECCV 2012, Part III, LNCS 7574, pp. 1–14, 2012.

invariant properties of shape. By performing an intrinsic regression analysis on the manifold of shapes, such factors are guaranteed to not effect the analysis.

In previous work, non-parametric kernel-based and spline-based methods have been extended to observations that lie on a Riemannian manifold with some success [2,3], but intrinsic parametric regression on Riemannian manifolds has received limited attention. Most recently, Fletcher [4] and Niethammer et al. [5] have each independently developed geodesic regression which generalizes the notion of linear regression to Riemannian manifolds.

The goal of the current work is to extend the geodesic model in order to accommodate more flexibility while remaining in the parametric setting. The increased flexibility introduced by the methods in this manuscript allow a better description of the variability in the data, and ultimately will allow more powerful statistical inference. Our work builds off that of Jupp & Kent [3], whose method for fitting parametric curves to the sphere involved intricate unwrapping and rolling processes. The work presented in this paper allows one to fit polynomial regression curves on a general Riemannian manifold, using intrinsic methods and avoiding the need for unwrapping and rolling. Since our model includes time-reparametrized geodesics as a special case, information about time dependence may be obtained from the regression without explicit modeling by examining the collinearity of the estimated parameters.

The generality of the Riemannian framework developed in this work allows this parametric regression technique to be applied in a wide variety of contexts besides shape analysis. Rigid and non-rigid motion can be modeled with polynomials in the Lie groups $\mathrm{SO}(3), \mathrm{SE}(3)$, or $\mathrm{Diff}(\Omega)$ [6,7]. Alternatively, the action of a Riemannian Lie group on a homogeneous space is often a convenient setting, such as in the currents framework [8], or the action of diffeomorphisms on images [9]. Additionally, there are applications in which the data are best modeled directly in a Riemannian manifold, such as modeling directional data on the sphere, or modeling motion via Grassmannian manifolds [10].

We demonstrate our algorithm in three studies of shape. By applying our algorithm to Bookstein's classical rat skull growth dataset [11], we show that we are able to obtain a parametric regression curve of similar quality to that produced by non-parametric methods [12]. We also demonstrate in a 2D corpus callosum aging study that, in addition to providing more flexibility in the traced path, our polynomial model provides information about the optimal parametrization of the time variable. Finally, the usefulness of shape regression for 3D shapes is demonstrated in an infant cerebellum growth study.

2 Methods

2.1 Preliminaries

Let (M, g) be a Riemannian manifold [13]. For each point $p \in M$, the metric g determines an inner product on the tangent space T_pM as well as a way to

differentiate vector fields with respect to one another. That derivative is referred to as the covariant derivative and is denoted by ∇. If $v, w$are smooth vector fields on M and $\gamma : [0, T] \to M$ is a smooth curve on M then the covariant derivative satisfies the product rule

$$\frac{d}{dt}\langle v(\gamma(t)), w(\gamma(t))\rangle = \langle \nabla_{\dot\gamma} v(\gamma(t)), w(\gamma(t))\rangle + \langle v(\gamma(t)), \nabla_{\dot\gamma} w(\gamma(t))\rangle \tag{1}$$

Geodesics $\gamma : [0, T] \to M$ are characterized by the conservation of kinetic energy along the curve:

$$\frac{d}{dt}\langle \dot\gamma, \dot\gamma\rangle = 0 = 2\langle \nabla_{\dot\gamma}\dot\gamma, \dot\gamma\rangle, \tag{2}$$

which is satisfied by the covariant differential equation (c.f. [13])

$$\nabla_{\dot\gamma}\dot\gamma = 0. \tag{3}$$

This differential equation, called the geodesic equation, uniquely determines geodesics parametrized by the initial conditions $(\gamma(0), \dot\gamma(0)) \in TM$. The mapping from the tangent space at p into the manifold is called the exponential map, $\mathrm{Exp}_p : T_pM \to M$. The exponential map is injective on a zero-centered ball B in T_pM of some non-zero radius. Thus for a point q in the image of B under Exp_p there exists a unique vector $v \in T_pM$ corresponding to a minimal length path under the exponential map from p to q. The mapping of such points q to their associated tangent vectors v at p is called the log map of q at p, denoted $v = \mathrm{Log}_p\, q$.

Given a curve $\gamma : [0, T] \to M$ we'll want to relate tangent vectors at different points along the curve. These relations are governed infinitesimally by the covariant derivative $\nabla_{\dot\gamma}$. A vector field v is parallel transported along γ if it satisfies the parallel transport equation, $\nabla_{\dot\gamma} v(\gamma(t)) = 0$. Notice that the geodesic equation is a special case of parallel transport, in which we require that the velocity is parallel along the curve itself.

2.2 Riemannian Polynomials

Given a vector field w along a curve γ, the covariant derivative of w gives us a way to define vector fields which are "constant" along γ, as parallel transported vectors. Geodesics are generalizations to the Riemannian manifold setting of curves in $\mathbb{R}^d$ with constant first derivative.

We will apply the covariant derivative repeatedly to examine curves with parallel higher derivatives. For instance, we refer to the vector field $\nabla_{\dot\gamma(t)}\dot\gamma(t)$ as the acceleration of the curve γ. Curves with parallel acceleration are generalizations of quadratic curves in $\mathbb{R}$ and satisfy the second order polynomial equation, $(\nabla_{\dot\gamma})^2\dot\gamma(t) = 0$. Extending this idea, a cubic polynomial is defined as a curve having parallel jerk (time derivative of acceleration), and so on. Generally, a kth order polynomial in M is defined as a curve $\gamma : [0, T] \to M$ satisfying

$$(\nabla_{\dot\gamma})^k\dot\gamma(t) = 0 \tag{4}$$

for all times $t \in [0, T]$ with initial conditions $\gamma(0)$ and $\dot{\gamma}^i(0), i = 1, \cdots, k$. As with polynomials in Euclidean space, a k^{th} order polynomial is parametrized by the $(k+1)$ initial conditions at $t = 0$.

Riemannian polynomials are alternatively described using the rolling maps of Jupp & Kent [3]. Suppose $\gamma(t)$ is a smooth curve in M. Then consider embedding M in m-dimensional Euclidean space and *rolling* M along a d-dimensional plane in $\mathbb{R}^m$ such that the plane is tangent to M at $\gamma(t)$ and the plane does not slip or twist with respect to M while rolling. The property that γ is a k-order polynomial is equivalent to the corresponding curve in $\mathbb{R}^d$ being a k-order polynomial in the conventional sense. For more information regarding this connection to Jupp & Kent's rolling maps, as well as a comparison to Noakes' cubic splines, the reader is referred to the literature of Leite & Krakowski [14].

The covariant differential equation governing the evolution of Riemannian polynomials is linearized in the same way that a Euclidean ordinary differential equation is linearized. Introducing vector fields $v_1(t), \ldots, v_k(t) \in T_{\gamma(t)}M$, we write the system of covariant differential equations as

$$\dot{\gamma}(t) = v_1(t) \tag{5}$$

$$\nabla_{\dot{\gamma}} v_i(t) = v_{i+1}(t), \quad i = 1, \ldots, k-1 \tag{6}$$

$$\nabla_{\dot{\gamma}} v_k(t) = 0. \tag{7}$$

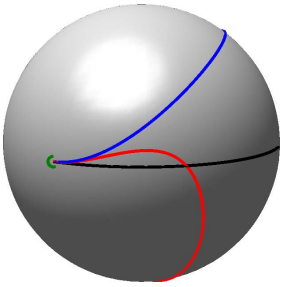

Fig. 1. Sample polynomial curves emanating from a common basepoint (green) on the sphere (black=geodesic, blue=quadratic, red=cubic)

The Riemannian polynomial equations cannot, in general, be solved in closed form, and must be integrated numerically. In order to discretize this system of covariant differential equations, we implement a covariant Euler integrator. At each step of the integrator, each vector is incremented within the tangent space at $\gamma(t)$ and the results are parallel transported infinitesimally along a geodesic from $\gamma(t)$ to $\gamma(t + \Delta t)$. The only ingredients necessary to integrate a polynomial forward in time are the exponential map and parallel transport on the manifold.

Figure 1 shows the result of integrating polynomials of order one, two, and three. The parameters, the initial velocity, acceleration, and jerk, were chosen a priori and a cubic polynomial was integrated to obtain the blue curve. Then the initial jerk was set to zero and the blue quadratic curve was integrated, followed by the black geodesic whose acceleration was also set to zero.

2.3 Estimation of Parameters for Regression

In order to regress polynomials against observed data $y_j \in M, j = 1, \ldots, N$ at known times $t_j \in \mathbb{R}, j = 1, \ldots, N$, we define the following constrained objective function

$$E_0(\gamma(0), v_1(0), \ldots, v_k(0)) = \frac{1}{N} \sum_{j=1}^{N} d(\gamma(t_j), y_j)^2 \tag{8}$$

subject to Eqs. 5–7. Note that in this expression d represents the geodesic distance: the minimum length of paths from the curve point $\gamma(t_j)$ to the data point y_j. The function E_0 is minimized in order to find the optimal initial conditions $\gamma(0), v_i(0), i = 1, \ldots, k$, which we will refer to as the parameters of our model.

In order to determine the optimal parameters $\gamma(0), v_i(0), i = 1, \ldots, k$, we introduce Lagrange multiplier vector fields $\lambda_i, i = 0, \ldots, k$, often called the adjoint variables, and define the unconstrained objective function

$$\begin{aligned} E(\gamma, \{v_i\}, \{\lambda_i\}) = \frac{1}{N} \sum_{j=1}^{N} d(\gamma(t_j), y_j)^2 + \int_0^T \langle \lambda_0(t), \dot{\gamma}(t) - v_1(t) \rangle dt \\ + \sum_{i=1}^{k-1} \int_0^T \langle \lambda_i(t), \nabla_{\dot{\gamma}} v_i(t) - v_{i+1}(t) \rangle dt + \int_0^T \langle \lambda_k(t), \nabla_{\dot{\gamma}} v_k(t) \rangle dt. \end{aligned} \tag{9}$$

As is standard practice, the optimality conditions for this equation are obtained by taking variations with respect to all arguments of E, integrating by parts when necessary. The resulting variations with respect the the adjoint variables yield the original dynamic constraints: the polynomial equations. Variations with respect to the primal variables gives rise to the following system of equations, termed the adjoint equations. The adjoint equations take the following form:

$$\nabla_{\dot{\gamma}} \lambda_i(t) = -\lambda_{i-1}(t) \quad i = 1, \ldots, k \tag{10}$$

$$\nabla_{\dot{\gamma}} \lambda_0(t) = -\frac{2}{N} \sum_{j=1}^{N} \delta(t - t_j) \operatorname{Log}_{\gamma} y_j - \sum_{i=1}^{k} R(v_i(t), \lambda_i(t)) v_1(t), \tag{11}$$

where R is the Riemannian curvature tensor and the Dirac δ functional indicates that the order zero adjoint variable takes on jump discontinuities at time points where data is present. The curvature tensor R satisfies the standard formula [13]

$$R(u, v)w = \nabla_u \nabla_v w - \nabla_v \nabla_u w - \nabla_{[u,v]} w, \tag{12}$$

and can be computed in closed form for many manifolds. Gradients of E with respect to initial and final conditions give rise to the terminal endpoint conditions for the adjoint variables, as well as expressions for the gradients with respect to the parameters $\gamma(0), v_i(0)$:

$$\lambda_i(T) = 0, \quad \delta_{\gamma(0)} E = -\lambda_0(0), \quad \delta_{v_i(0)} E = -\lambda_i(0) \qquad i = 0, \cdots, k \tag{13}$$

In order to determine the value of the adjoint vector fields at $t = 0$, and thus the gradients of the functional E_0, the adjoint variables are initialized to zero at time T, then Eq. 11 is evolved backward in time to $t = 0$.

Given the gradients with respect to the parameters, a simple steepest descent algorithm is used to optimize the functional. At each iteration, $\gamma(0)$ is updated using the exponential map and the vectors $v_i(0)$ are updated via parallel translation:

$$\gamma(0) \mapsto \mathrm{Exp}_{\gamma(0)}(-\epsilon \delta_{\gamma(0)} E) = \mathrm{Exp}_{\gamma(0)}(\epsilon \lambda_0(0)) \tag{14}$$

$$v_i(0) \mapsto \mathrm{ParTrans}_{\gamma(0)}(v_i(0) + \epsilon \lambda_i(0), \epsilon \lambda_0(0)), \tag{15}$$

where $\mathrm{ParTrans}_p(u, v)$ denotes the parallel translation of a vector u at p for unit time along a geodesic in the direction of v.

Note that in the special case of a zero-order polynomial ($k = 0$), the only gradient λ_0 is simply the mean of the log map vectors at the current estimate of the Fréchet mean. So this method generalizes the common method of Fréchet averaging on manifolds [15]. The curvature term, in the case $k = 1$, indicates that λ_1 is a sum of Jacobi fields. So this approach subsumes geodesic regression as presented by Fletcher [4]. For higher order polynomials, the adjoint equations represent a generalized Jacobi field.

In practice, it is often not necessary to explicitly compute the curvature terms. In the case that the manifold M embeds into a Hilbert space, the extrinsic adjoint equations can be computed by taking variations in the ambient space, using standard methods. Such an approach gives rise to the regression algorithm found in Niethammer et al. [5], for example.

2.4 Time Reparametrization

Geodesic curves propagate at a constant speed, a result of their extremal action property. However, using higher-order polynomials it is possible to generate a polynomial curve whose path (the image of the mapping γ) matches that of a geodesic, but whose time-dependence has been reparametrized. If the parameters consist of collinear v_i, then this is the case. In this way, polynomials provide flexibility not only in the class of paths that are possible, but in the time dependence of the curves traversing those paths. Regression models could even be implemented in which the operator wishes to estimate geodesic paths, but is unsure of parametrization, and so enforces the estimated parameters to be collinear.

2.5 Coefficient of Determination for Regression in Metric Spaces

It will be useful to define a statistic which will indicate how well our model fits a set of data. As in [4], we compute the coefficient of determination of our regression polynomial $\gamma(t)$, denoted R^2. The first step to computing R^2 is to

compute the variance of the data. The natural choice of total variance statistic is the Fréchet variance, defined by

$$\mathrm{Var}\{y_j\} = \frac{1}{N} \min_{\bar{y} \in M} \sum_{j=1}^{N} d(\bar{y}, y_j)^2. \tag{16}$$

Note that the Fréchet mean $\bar{y}$ itself is the 0-th order polynomial regression against the data $\{y_j\}$ and the variance is the value of the cost function E_0 at that point. We define the sum of squared error for a curve γ as the value $E_0(\gamma)$:

$$SSE = \frac{1}{N} \sum_{j=1}^{N} d(\gamma(t_j), y_j)^2. \tag{17}$$

Then the coefficient of determination is defined as

$$R^2 = 1 - \frac{SSE}{\mathrm{Var}\{y_j\}}. \tag{18}$$

Any regression polynomial will have $SSE < \mathrm{Var}\{y_j\}$ and R^2 will have a value between 0 and 1, with 1 indicating a perfect fit and 0 indicating that the curve γ provides no better fit than does the Fréchet mean.

3 Examples

3.1 Lie Groups

A special case arises when the manifold M is a Lie group and the metric g is left-invariant. In this case, left-translation by $\gamma(t)$ and $\gamma^{-1}(t)$, denoted $L_\gamma, L_{\gamma^{-1}}$, allow one to isometrically represent tangent vectors at $\gamma(t)$ as tangent vectors at the identity, using the pushforward $L_{\gamma^{-1}*}$ [6]. The tangent space at the identity is isomorphically identified with the Lie algebra of the group, $\mathfrak{g}$. The algebraic operation in $\mathfrak{g}$ is called the infinitesimal adjoint action $\mathrm{ad} : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$. Fixing X, the adjoint of the operator ad_X satisfies $\langle \mathrm{ad}_X Y, Z \rangle = \langle Y, \mathrm{ad}_X^\dagger Z \rangle$. We will refer to the operator $\mathrm{ad}_X^\dagger$ as the adjoint-transpose action of X.

Suppose X is a vector field along the curve γ and $X_c = L_{\gamma^{-1}*} X, \omega_c = L_{\gamma^{-1}*} \dot{\gamma} \in \mathfrak{g}$ is its representative at the identity. Similarly, $\omega_c = L_{\gamma^{-1}*} \dot{\gamma}$ is the representative of the curve's velocity in the Lie algebra. Then the covariant derivative of X along γ evolves in the Lie algebra via

$$L_{\gamma^{-1}*} \nabla_{\dot{\gamma}} X = \dot{X}_c - \frac{1}{2} \left(\mathrm{ad}_{\omega_c}^\dagger X_c + \mathrm{ad}_{X_c}^\dagger \omega_c - \mathrm{ad}_{\omega_c} X_c \right). \tag{19}$$

In the special case when $X = \dot{\gamma}$ so that $X_c = \omega_c$, if we set this covariant derivative equal to zero we have the geodesic equation:

$$\dot{\omega_c} = \mathrm{ad}_{\omega_c}^\dagger \omega_c, \tag{20}$$

which in this form is often refered to as the Euler-Poincaré equation [6,16]. Also note that if the metric is both left- and right-invariant, then the adjoint-transpose action is alternating and we can simplify the covariant derivative further. Parallel transport in such case requires solution of the equation $\left(\frac{d}{dt} + \mathrm{ad}_{\omega_c}\right) X = 0$, which can be integrated using Euler integration or similar time-stepping algorithms, so long as the adjoint action can be easily computed. Because of the simplified covariant derivative formula for a bi-invariant metric, the curvature also takes a simple form [13]: $R(X,Y)Z = \frac{1}{4}\,\mathrm{ad}_Z\,\mathrm{ad}_X\,Y$.

The Lie Group SO(3). In order to illustrate the algorithm in a Lie group, we consider the Lie group SO(3) of orthogonal 3-by-3 real matrices with determinant one. We review the basics of SO(3) here, but the reader is encouraged to consult [6] for a full treatment of SO(3) and derivation of the Euler-Poincaré equation there.

The Lie algebra for the rotation group is $\mathfrak{so}(3)$, the set of all skew-symmetric 3-by-3 real matrices. Such matrices can be identified with vectors in $\mathbb{R}^3$ by the identification

$$x = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \mapsto \sigma(x) = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix} \tag{21}$$

under which the infinitesimal adjoint action takes the convenient form of the cross-product of vectors:

$$\mathrm{ad}_x\, y = x \times y = \sigma(x)y \mapsto \sigma(x \times y) = [\sigma(x), \sigma(y)]. \tag{22}$$

With the $\mathbb{R}^3$ representation of vectors in $\mathfrak{so}(3)$, a natural choice of bi-invariant metric is available: $\langle x, y\rangle = x^T y$. With this metric the adjoint-transpose action is $\mathrm{ad}_x^\dagger\, y = -x \times y$. If we plug this into the general Lie group formula we obtain the geodesic equation in the Lie algebra:

$$\dot{\omega_c} = -(\omega_c \times \omega_c) = 0, \tag{23}$$

reflecting the fact that angular momentum is conserved under geodesics in SO(3). The exponential map is simply the matrix exponential of ω_c, which can be computed in closed form using Rodrigues' rotation formula [6]. To compute a polynomial curve, the matrix v is converted to a skew-symmetric matrix by multiplication by $\gamma(t)^{-1}$, which translates v back into the Lie algebra. The resulting skew-symmetric matrix is converted to a vector $\omega_c(0)$. The polynomial equation is then integrated to find ω_c at future times. The point $\gamma(t)$ then evolves along the exponential map by time-stepping, where at each step ω_c is converted back into a skew-symmetric matrix W, multiplied by Δt and exponentiated to obtain a rotation matrix. The Euler integration step is then given by left-multiplication by the resulting matrix.

Plugging into the above expressions, the curvature of SO(3) is

$$R(u,v)w = \frac{1}{4} w \times u \times v. \tag{24}$$

and the parallel transport equation is given by $\dot{u_c} = -\frac{1}{2}\omega_c \times u_c$.

3.2 The n-Dimensional Sphere

In the case that M is an n-dimensional sphere, $S^n = \{p \in \mathbb{R}^{n+1} : \|p\| = 1\}$, a natural metric exists corresponding to that of $\mathbb{R}^{n+1}$ and to the bi-invariant metric on $\mathrm{SO}(n+1)$, under which S^n is a homogeneous space. Geodesics are great circles and the exponential map, log map, and Riemannian curvature are given by

$$\mathrm{Exp}_p\, v = \cos\theta \cdot p + \sin\theta \frac{v}{\|v\|}, \quad \theta = \|v\| \tag{25}$$

$$\mathrm{Log}_p\, q = \theta \frac{q - (p^T q)p}{\|q - (p^T q)p\|}, \quad \theta = \cos^{-1}(p^T q). \tag{26}$$

$$R(u,v)w = (u^T w)v - (v^T w)u. \tag{27}$$

Parallel translation of a vector w along the exponential map of a vector v is performed as follows. The vector w is decomposed into a vector parallel to v, which we denote w^v, and a vector orthogonal to v, $w^\perp$. $w^\perp$ is left unchanged by parallel transport along $\mathrm{Exp}_p\, v$, while w^v transforms in a similar way as v:

$$w^v \mapsto \left(-p\sin\theta + \frac{v}{\|v\|}\cos\theta\right)\frac{v^T}{\|v\|}w^v. \tag{28}$$

Using this exponential map and parallel transport, a polynomial integrator was implemented. Figure 1 shows the result of integrating polynomials of order one, two, and three, using the equations from this section, for the special case of a two-dimensional sphere.

3.3 Kendall Shape Space

We will briefly describe Kendall's shape space of m-landmark point sets in $\mathbb{R}^d$, denoted Σ_d^m. For a more complete overview of Kendall's shape space, including computation of the curvature tensor, the reader is encouraged to consult the work of Kendall and Le [17,18].

Let $x = (x_i)_{i=1,\ldots,m}, x_i \in \mathbb{R}^d$ be a labelled point set in $\mathbb{R}^d$. Translate the point set so that the centroid resides at 0, removing any effects due to translations. Another degree of freedom due to uniform scaling is removed by requiring that $\sum_{i=1}^m \|x_i\|^2 = 1$. After this standardization, x is a point in the sphere S^{md-1}, commonly referred to as *preshape space* in this context. The Kendall shape space Σ_d^m is obtained by taking the quotient of the preshape space by the action of the rotation group $\mathrm{SO}(d)$ on point sets. In practice, points in the quotient (referred to as *shapes*) are represented by a member of their equivalence class in preshape space. The work of O'Neill [19] characterizes the link in geometry between the shape and preshape spaces.

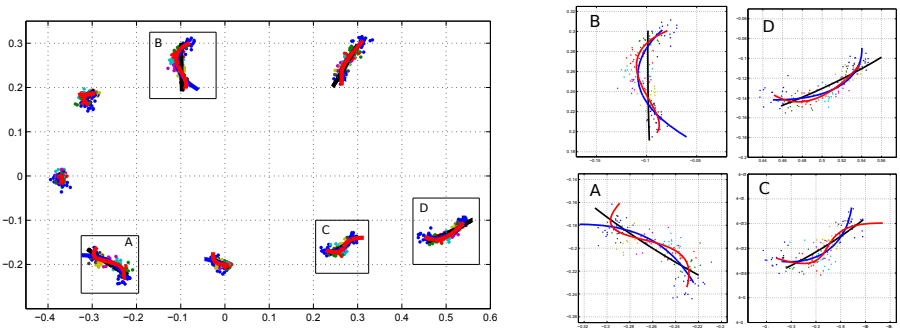

Fig. 2. Bookstein rat calivarium data after uniform scaling and Procrustes alignment. The colors of dots indicate age group while the colors of lines indicate order of polynomial used for the regression (black=geodesic, blue=quadratic, red=cubic). Zoomed views of individual rectangles are shown at bottom. Note that the axes are arbitrary, due to scale-invariance of Kendall shape space.

For a general shape space Σ_d^m, the exponential map and parallel translation are performed using representatives in preshape space, S^{md-1}. For $d > 2$, this must be done in a time-stepping algorithm, in which at each time step an infinitesimal spherical parallel transport is performed, followed by the horizontal projection. The resulting algorithm can be used to compute the exponential map as well. Computation of the log map is less trivial, as it requires an iterative optimization routine. However, a special case arises in the case when $d = 2$. In this case the exponential map, parallel transport and log map have closed form. The reader is encouraged to consult [4] for more details about the two-dimensional case. With exponential map, log map, and parallel transport implemented as described here, one can perform polynomial regression on Kendall shape space in any dimension. We demonstrate such regression now on examples in both two and three dimensions.

4 Results

4.1 Rat Calivarium Growth

The first dataset we consider was first analyzed by Bookstein [11]. The data is available for download at `http://life.bio.sunysb.edu/morph/data/datasets.html` and consists of $m = 8$ landmarks on a midsagittal section of rat calivaria (upper skulls). Landmark positions are available for 18 rats and at 8 ages apiece. Riemannian polynomials of orders $k = 0, 1, 2, 3$ were fit to this data. The resulting curves are shown in Fig. 2. Clearly the quadratic and cubic curves differ from that of the geodesic regression. The R^2 values agree with this qualitative difference: the geodesic regression has $R^2 = 0.79$, while the quadratic and cubic regressions both have R^2 values of 0.85 and 0.87, respectively (see Table 1). While this shows that there is a clear improvement in the fit due to increasing k from one to two, it

also shows that little is gained by further increasing the order of the polynomial. Qualitatively, Fig. 2 shows that the slight increase in R^2 obtained by moving from a quadratic to cubic model corresponds to a marked difference in the curves, possibly indicating that the cubic curve is overfitting the data.

Corpus Callosum Aging. The corpus callosum is the major white matter bundle connecting the two hemispheres of the brain. In order to investigate shape change of the corpus callosum during normal aging, polynomial regression was performed on a collection of data from the OASIS brain database (`www.oasis-brains.org`). Magnetic resonance imaging (MRI) scans from 32 normal subjects with ages between 19 and 90 years were obtained from the database. A midsagittal slice was extracted from each volumetric image and segmented using the ITK-SNAP program (`www.itksnap.org`). Sets of 64 landmarks were optimized on the contours of these segmentations using the ShapeWorks program [20] (`www.sci.utah.edu/software.html`). The algorithm generates samplings of each shape boundary with optimal correspondences among the populaton.

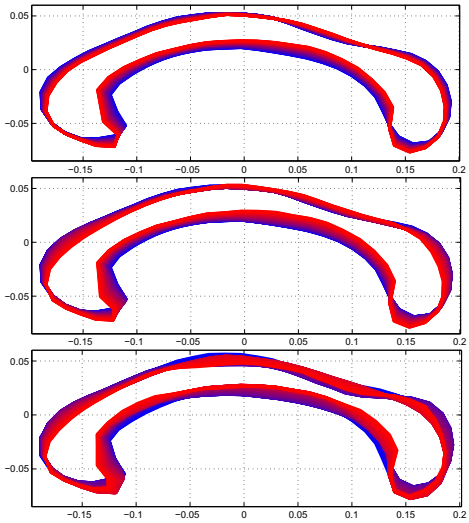

Fig. 3. Geodesic (top, $R^2 = 0.12$) quadratic (middle, $R^2 = 0.13$) and cubic (bottom, $R^2 = 0.21$) regression for corpus callosum dataset. Color represents age, with blue indicating youth (age 19) and red indicating old age (age 90).

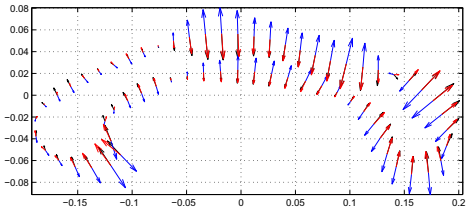

Fig. 4. Parameters for cubic regression of corpus callosa. The velocity (black), acceleration (blue) and jerk (red) are nearly collinear.

Regression results for geodesic, quadratic, and cubic regression are shown in Fig. 3. At first glance the results appear similar for the three different models, since the motion envelopes show close agreement. However, as shown in Table 1, the R^2 values show an improvement from geodesic to quadratic (from 0.12 to 0.13) and from quadratic to cubic (from 0.13 to 0.21). Inspection of the estimated parameters, shown in Fig. 4 reveals that the tangent vectors appear to be rather collinear. For the reasons stated in Sec. 2.4,

this suggests that the differences can essentially be described as reparametrization of the time variable, which is only accommodated in the higher order polynomial models.

Infant Cerebellum Growth. To test the performance of the 3D shape space algorithm, polynomials were fit to a set of ten cerebellum shapes extracted from structural MRI data. The data were obtained in clinical studies of subjects with ages ranging from neonatal to 4 years old. A three-dimensional 1.5T, T1-weighted MRI was obtained for each patient, which was then segmented to obtain a volume segmentation of the cerebellum. ShapeWorks was then used to find correspondences among the ten shapes, using 256 landmarks for each shape. A sample point-set is shown in Fig. 5.

Table 1. R^2 for regression of three datasets and three orders of polynomial (k)

k	Rat Calivarium	Corpus Callosum	Cerebellum
1	.79	0.12	0.082
2	.85	0.13	0.17
3	.87	0.21	0.19

Polynomial regression was performed on the cerebellum using geodesic, quadratic, and cubic curves. The R^2 value for the geodesic regression was 0.082. The quadratic regression fit the data much better, improving the R^2 value to 0.170. The cubic polynomial showed marginal improvement over quadratic, resulting in an of R^2 value of 0.192 (see Table 1. Figure 6 shows the estimated parameters for the cubic polynomial regression. Notice that the hemispheres of the cerebellum show smaller tangent vectors than those on the vermis (the ridge dividing the two hemispheres). This indicates that the major modes of shape change due to age are likely related to changes in the vermis.

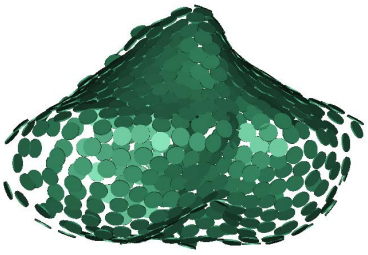

Fig. 5. Typical cerebellum data used for regression.

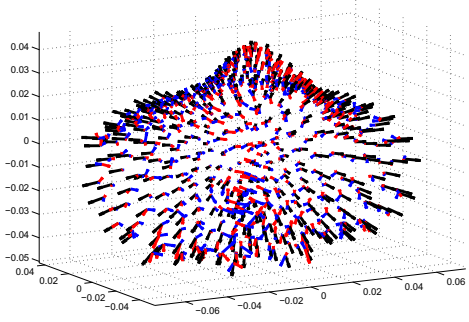

Fig. 6. Estimated parameters for cubic regression on cerebellum data. Velocity is shown in black, with acceleration in blue, and jerk shown in red.

5 Conclusion

Modern regression analysis consists of two classes of techniques: parametric and nonparametric. Nonparametric approaches to curve-fitting include kernel-based methods and splines, each of which have been explored in the setting of Riemannian manifolds [21,22,23,24,25,26]. Such approaches offer increased flexibility over parametric models and the ability to accurately model observed data, at the expense of efficiency and precision. The trade-offs between parametric and nonparametric regression are well-established [27] and beyond the scope of this work. A large sample size is required for cross-validation to estimate the optimal smoothing parameter in a spline-based approach or to determine the kernel bandwidth in a kernel smoothing method and can be computationally prohibitive. Compounding this problem, since the number of parameters is tied to sample size, these nonparametric fits will vary greatly due to the curse of dimensionality. Parametric analyses enable robust fitting with sample sizes that are more feasible in biomedical applications. This tradeoff between accuracy and precision must be considered on a per-application basis when selecting a regression model.

In this work, we presented for the first time a methodology for parametric regression on Riemannian manifolds, with the parameters being the initial conditions of the curve. Our results show that the polynomial model allows for parametric analysis on manifolds while providing more flexibility and accuracy than geodesics. For the classical Bookstein rat skull dataset for example, we are able to provide a rather high R^2 coefficient of 0.84 with only a quadratic model.

Although we have presented results for landmark shape data, the framework is general and can be applied to different data and modes of variation. For instance, landmark data without direct correspondences can be modeled in the currents framework [8], and image or surface data can be fit using the popular diffeomorphic matching framework [7,9,5].

Acknowledgements. This work was made possible in part by NIH/NIBIB Grant 5R01EB007688, NIH/NCRR Grant 2P41 RR0112553-12, and by NSF CAREER Grant 1054057. We also acknowledge support from the NIMH Silvio Conte Center for Neuroscience of Mental Disorders MH064065 and the BRP grant R01 NS055754-01-02.

References

1. Dryden, I.L., Mardia, K.V.: Statistical Shape Analysis. Wiley (1998)
2. Davis, B.C., Fletcher, P.T., Bullitt, E., Joshi, S.C.: Population shape regression from random design data. Int. J. Comp. Vis. 90, 255–266 (2010)
3. Jupp, P.E., Kent, J.T.: Fitting smooth paths to spherical data. Appl. Statist. 36, 34–46 (1987)
4. Fletcher, P.T.: Geodesic regression on Riemannian manifolds. In: International Workshop on Mathematical Foundations of Computational Anatomy, MFCA (2011)

5. Niethammer, M., Huang, Y., Vialard, F.-X.: Geodesic Regression for Image Time-Series. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part II. LNCS, vol. 6892, pp. 655–662. Springer, Heidelberg (2011)
6. Arnol'd, V.I.: Mathematical Methods of Classical Mechanics, 2nd edn. Springer (1989)
7. Cootes, T.F., Twining, C.J., Taylor, C.J.: Diffeomorphic statistical shape models. In: BMVC (2004)
8. Vaillant, M., Glaunés, J.: Surface matching via currents. In: IPMI (2005)
9. Miller, M.I., Trouvé, A., Younes, L.: Geodesic shooting for computational anatomy. Journal of Mathematical Imaging and Vision 24, 209–228 (2006)
10. Turaga, P., Veeraraghavan, A., Srivastava, A., Chellappa, R.: Statistical computations on Grassmann and Stiefel manifolds for image and video-based recognition 33, 2273–2286 (2011)
11. Bookstein, F.L.: Morphometric Tools for Landmark Data: Geometry and Biology. Cambridge Univ. Press (1991)
12. Kent, J.T., Mardia, K.V., Morris, R.J., Aykroyd, R.G.: Functional models of growth for landmark data. In: Proceedings in Functional and Spatial Data Analysis, pp. 109–115 (2001)
13. do Carmo, M.P.: Riemannian Geometry, 1st edn. Birkhäuser, Boston (1992)
14. Leite, F.S., Krakowski, K.A.: Covariant differentiation under rolling maps. Centro de Matemática da Universidade de Coimbra (2008) (preprint)
15. Fletcher, P.T., Liu, C., Pizer, S.M., Joshi, S.C.: Principal geodesic analysis for the study of nonlinear statistics of shape. IEEE Trans. Med. Imag. 23, 995–1005 (2004)
16. Holm, D.D., Marsden, J.E., Ratiu, T.S.: The Euler-Poincaré equations and semidirect products with applications to continuum theories. Adv. in Math. 137, 1–81 (1998)
17. Kendall, D.G.: A survey of the statistical theory of shape. Statistical Science 4, 87–99 (1989)
18. Le, H., Kendall, D.G.: The Riemannian structure of Euclidean shape spaces: A novel environment for statistics. Ann. Statist. 21, 1225–1271 (1993)
19. O'Neill, B.: The fundamental equations of a submersion. Michigan Math J. 13, 459–469 (1966)
20. Cates, J., Fletcher, P.T., Styner, M., Shenton, M., Whitaker, R.: Shape modeling and analysis with entropy-based particle systems. In: IPMI (2007)
21. Davis, B.C., Fletcher, P.T., Bullitt, E., Joshi, S.: Population shape regression from random design data. In: ICCV (2007)
22. Noakes, L., Heinzinger, G., Paden, B.: Cubic splines on curved surfaces. IMA J. Math. Control Inform. 6, 465–473 (1989)
23. Giambò, R., Giannoni, F., Piccione, P.: An analytical theory for Riemannian cubic polynomials. IMA J. Math. Control Inform. 19, 445–460 (2002)
24. Machado, L., Leite, F.S.: Fitting smooth paths on Riemannian manifolds. Int. J. App. Math. Stat. 4, 25–53 (2006)
25. Samir, C., Absil, P.A., Srivastava, A., Klassen, E.: A gradient-descent method for curve fitting on Riemannian manifolds. Found. Comp. Math. 12, 49–73 (2010)
26. Su, J., Dryden, I.L., Klassen, E., Le, H., Srivastava, A.: Fitting smoothing splines to time-indexed, noisy points on nonlinear manifolds. Image and Vision Computing 30, 428–442 (2012)
27. Moussa, M.A.A., Cheema, M.Y.: Non-parametric regression in curve fitting. The Statistician 41, 209–225 (1992)