In proceedings of Computer Vision Approaches
to Medical Image Analysis 2004 (CVAMIA)

Principal Geodesic Analysis on Symmetric Spaces: Statistics of Diffusion Tensors

P. Thomas Fletcher and Sarang Joshi

Medical Image Display and Analysis Group,
University of North Carolina at Chapel Hill
fletcher@cs.unc.edu

Abstract. Diffusion tensor magnetic resonance imaging (DT-MRI) is emerging as an important tool in medical image analysis of the brain. However, relatively little work has been done on producing statistics of diffusion tensors. A main difficulty is that the space of diffusion tensors, i.e., the space of symmetric, positive-definite matrices, does not form a vector space. Therefore, standard linear statistical techniques do not apply. We show that the space of diffusion tensors is a type of curved manifold known as a Riemannian symmetric space. We then develop methods for producing statistics, namely averages and modes of variation, in this space. In our previous work we introduced principal geodesic analysis, a generalization of principal component analysis, to compute the modes of variation of data in Lie groups. In this work we expand the method of principal geodesic analysis to symmetric spaces and apply it to the computation of the variability of diffusion tensor data. We expect that these methods will be useful in the registration of diffusion tensor images, the production of statistical atlases from diffusion tensor data, and the quantification of the anatomical variability caused by disease.

1 Introduction

Diffusion tensor magnetic resonance imaging (DT-MRI) [2] produces a 3D diffusion tensor, i.e., a 3×3, symmetric, positive-definite matrix, at each voxel of an imaging volume. This tensor is the covariance in a Brownian motion model of the diffusion of water at that voxel. In brain imaging DT-MRI is used to track the white matter fibers, which demonstrate higher diffusivity of water in the direction of the fiber. The aim of this paper is to provide new methods for the statistical analysis of diffusion tensors.

Diffusion tensor imaging has shown promise in clinical studies of brain pathologies, such as multiple sclerosis and stroke, and in the study of brain connectivity [4]. Several authors have addressed the problem of estimation and smoothing within a DT image [6, 7, 14]. Further insights might be had from the use of diffusion tensor imaging in intersubject studies. Statistical brain atlases have been used in the case of scalar images to quantify anatomical variability across patients. However, relatively little work has been done towards constructing statistical brain atlases from diffusion tensor images. Alexander *et al.* [1] describe a method for the registration of multiple DT images into a common coordinate frame, however, they do not include a statistical analysis of the diffusion tensor data. Previous attempts [3, 12] at statistical analysis of diffusion tensors within a DT image are based on a Gaussian model of the linear tensor coefficients. In this paper we demonstrate that the space of diffusion tensors is more naturally described as a Riemannian symmetric space, rather than a linear space. In our previous work we introduced *principal geodesic analysis* (PGA) as an analog of principal component analysis for studying the statistical variability of Lie group data. Extending these ideas to symmetric spaces, we develop new methods for computing averages and describing the variability of diffusion tensor data. We show that these statistics preserve natural properties of the diffusion tensors, most importantly the positive-definiteness, that are not preserved by linear statistics. The framework presented in this paper thus provides the statistical methods needed for constructing statistical atlases of diffusion tensor images.

2 The Space of Diffusion Tensors

Recall that a real $n \times n$ matrix A is symmetric if $A = A^T$ and positive-definite if $x^T Ax > 0$ for all nonzero $x \in \mathbb{R}^n$. We denote the space of all $n \times n$ symmetric, positive-definite matrices as $P(n)$. The tensors in DT-MRI are thus elements of $P(3)$. The space $P(n)$ forms a convex subset of $\mathbb{R}^{n^2}$. One can define a linear average of N positive-definite, symmetric matrices $A_1, \ldots, A_N$ as $\mu = \frac{1}{N} \sum_{i=1}^{N} A_i$. This definition minimizes the Euclidean metric on $\mathbb{R}^{n^2}$. Since $P(n)$ is convex, μ is lies within $P(n)$, however, linear averages do not interpolate natural properties. The linear average of matrices of the same determinant can result in a matrix with a larger determinant. Second order statistics are even more problematic. The standard principal component analysis is invalid because the straight lines defined by the modes of variation do not stay within the space $P(n)$. In other words, linear PCA does not preserve the positive-definiteness of diffusion tensors. The reason for such difficulties is that space $P(n)$, although a subset of a vector space, is not a vector space, e.g., the negation of a positive-definite matrix is not positive-definite.

In this paper we derive a more natural metric on the space of diffusion tensors, $P(n)$, by viewing it not simply as a subset of $\mathbb{R}^{n^2}$, but rather as a Riemannian symmetric space. Following Fréchet [9], we define the average as the minimum mean squared error estimator under this metric. We develop the method of principal geodesic analysis to describe the variability of diffusion tensor data. Principal geodesic analysis is the generalization of principal component analysis to manifolds. In this framework the modes of variation are represented as flows along geodesic curves, i.e., shortest paths under the Riemannian metric. These geodesic curves, unlike the straight lines of $\mathbb{R}^{n^2}$, are completely contained within $P(n)$, that is, they preserve the positive-definiteness. Principal component analysis generates lower-dimensional subspaces that maximize the projected variance of the data. Thus the development of principal geodesic analysis requires that we generalize the concepts of variance and projection onto lower-dimensional subspaces for data in symmetric spaces.

To illustrate these issues, consider the space $P(2)$, the 2×2 symmetric, positive-definite matrices. A matrix $A \in P(2)$ is of the form

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \quad ac - b^2 > 0, \quad a > 0.$$

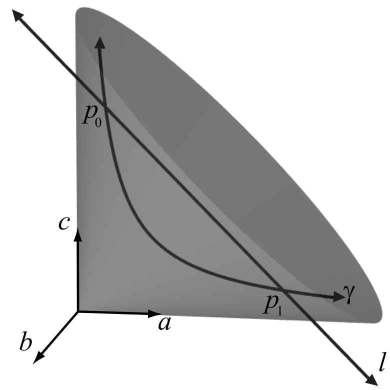

Fig. 1. The space $P(2)$, showing the geodesic γ and the straight line l between the two points p_0 and p_1.

If we consider the matrix A as a point $(a, b, c) \in \mathbb{R}^3$, then the above conditions describe the interior of a cone as shown in Fig. 1. The two labelled points are $p_0 = (1, 0, 7), p_1 = (7, 0, 1)$. The straight line l between the two points does not remain contained within the space $P(2)$. The curve γ is the geodesic between the two points when $P(2)$ is considered as a Riemannian symmetric space. This geodesic lies completely within $P(2)$. We chose $P(2)$ as an example since it can be easily visualized, but the same phenomenon occurs for general $P(n)$, i.e., $n > 2$.

3 The Geometry of $P(n)$

In this section we show that the space of diffusion tensors, $P(n)$, can be formulated as a Riemannian symmetric space. This leads to equations for computing geodesics that will be essential in defining the statistical methods for diffusion tensors. The differential geometry of diffusion tensors has also been used in [6], where the diffusion tensor smoothing was constrained along geodesic curves. A more thorough treatment of symmetric spaces can be found in [5, 10].

A *symmetric space* is a connected Riemannian manifold M such that for each $x \in M$ there is an isometry σ_x which (1) is involutive, i.e., $\sigma_x^2 = \mathrm{id}$, and (2) has x as an isolated fixed point, that is, there is a neighborhood U of x where σ_x leaves only x fixed. It can be shown that σ_x is the map that reverses all geodesics through the point x. Riemannian symmetric spaces, and the methods for computing geodesics and distances on them, arise naturally from Lie group actions on manifolds.

3.1 Lie Group Actions

A *Lie group* is an algebraic group G that also forms a differentiable manifold, where the two group operations, multiplication and inversion, are smooth mappings. Many common geometric transformations of Euclidean space form Lie groups. For example, rotations, translations, and affine transformations of $\mathbb{R}^n$ all form Lie groups. More generally, Lie groups can be used to describe transformations of smooth manifolds.

Given a manifold M and a Lie group G, a *smooth group action* of G on M, or *smooth G-action* on M, is a smooth mapping $\phi : G \times M \to M$ such that for all $g, h \in G$, and all $x \in M$ we have $\phi(e, x) = x$, and $\phi(g, \phi(h, x)) = \phi(gh, x)$, where e is the identity element of G. Consider the Lie group of all $n \times n$ real matrices with positive determinant, denoted $GL^+(n)$. This group acts on $P(n)$ via

$$
\begin{aligned}
\phi &: GL^+(n) \times P(n) \to P(n) \\
\phi(g, p) &= gpg^T. \qquad (1)
\end{aligned}
$$

The *orbit* under ϕ of a point $x \in M$ is defined as $G(x) = \{\phi(g, x) : g \in G\}$. In the case that M consists of a single orbit, we call M a *homogeneous space* and say that the G-action is *transitive*. The space $P(n)$ is a homogeneous space, as is easy to derive from the fact that any matrix $p \in P(n)$ can be decomposed as $p = gg^T = \phi(g, I_n)$, where $g \in GL^+(n)$ and I_n is the $n \times n$ identity matrix. The *isotropy subgroup* of x is defined as $G_x = \{g \in G : \phi(g, x) = x\}$, i.e., G_x is the subgroup of G that leaves the point x fixed. For $P(n)$ the isotropy subgroup of I_n is $SO(n) = \{g \in GL^+(n) : \phi(g, I_n) = gg^T = I_n\}$, i.e., the space of $n \times n$ rotation matrices.

Let H be a closed Lie subgroup of the Lie group G. Then the *left coset* of an element $g \in G$ is defined as $gH = \{gh : h \in H\}$. The space of all such cosets is denoted G/H and is a smooth manifold. There is a natural bijection $G(x) \cong G/G_x$ given by the mapping $g \cdot x \mapsto gG_x$. Therefore, we can consider the space of diffusion tensors, $P(n)$, as the coset space $GL^+(n)/SO(n)$. An intuitive way to view this is to think of the polar decomposition, which decomposes a matrix $g \in GL^+(n)$ as $g = pu$, where $p \in P(n)$ and $u \in SO(n)$. Thus, the diffusion tensor space $P(n) \cong GL^+(n)/SO(n)$ comes from "dividing out" the rotational component in the polar decomposition of $GL^+(n)$.

3.2 Invariant Metrics

A *Riemannian metric* on a manifold M smoothly assigns to each point $x \in M$ an inner product $\langle \cdot, \cdot \rangle_x$ on T_xM, the tangent space to M at x. If ϕ is a smooth G-action on M, a metric on M is called *G-invariant* if for each $g \in G$ the map $\phi_g : x \mapsto \phi(g, x)$ is an isometry, i.e., ϕ_g preserves distances on M. The space of diffusion tensors, $P(n)$, has a metric that is invariant under the $GL^+(n)$ action, which follows from the fact that the isotropy subgroup $SO(n)$ is connected and compact (see [5], Theorem 9.1).

The tangent space of $P(n)$ at the identity matrix can be identified with the space of $n \times n$ symmetric matrices, $\mathrm{Sym}(n)$. Since the group action $\phi_g : s \mapsto gsg^T$ is linear, its derivative map, denoted $d\phi_g$, is given by $d\phi_g(X) = gXg^T$. If $X \in \mathrm{Sym}(n)$, it is easy to see that $d\phi_g(X)$ is again a symmetric matrix. Thus the tangent space at any point $p \in P(n)$ is also equivalent to $\mathrm{Sym}(n)$. If $X, Y \in \mathrm{Sym}(n)$ represent two tangent vectors at $p \in P(n)$, where $p = gg^T, g \in GL^+(n)$, then the Riemannian metric at p is given by the inner product

$$\langle X, Y \rangle_p = \mathrm{tr}(g^{-1}Xp^{-1}Y(g^{-1})^T).$$

Finally, the mapping $\sigma_{I_n}(p) = p^{-1}$ is an isometry that reverses geodesics of $P(n)$ at the identity, and this turns $P(n)$ into a symmetric space.

3.3 Computing Geodesics

Geodesics on a symmetric space are easily derived via the group action (see [10] for details). Let p be a point on $P(n)$ and X a tangent vector at p. There is a unique geodesic, γ, with initial point $\gamma(0) = p$ and tangent vector $\gamma'(0) = X$. To derive an equation for such a geodesic, we begin with the special case where the initial point p is the $n \times n$ identity matrix, I_n, and the tangent vector X is diagonal. Then the geodesic is given by

$$\gamma(t) = \exp(tX),$$

where $\exp$ is the matrix exponential map given by the infinite series

$$\exp(X) = \sum_{k=0}^{\infty} \frac{1}{k!} X^k.$$

For the diagonal matrix X with entries x_i, the matrix exponential is simply the diagonal matrix with entries e^{x_i}.

Now for the general case consider the geodesic γ starting at an arbitrary point $p \in P(n)$ with arbitrary tangent vector $X \in \mathrm{Sym}(n)$. We will use the group action to map this configuration into the special case described above, i.e., with initial point at the identity and a diagonal tangent vector. Since the group action is an isometry, geodesics and distances are preserved. Let $p = gg^T$, where $g \in GL^+(n)$. Then the action $\phi_{g^{-1}}$ maps p to I_n. The tangent vector is mapped via the corresponding tangent map to $Y = d\phi_{g^{-1}}(X) = g^{-1}X(g^{-1})^T$. Now we may write $Y = v\Sigma v^T$, where v is a rotation matrix and Σ is diagonal. The group action $\phi_{v^{-1}}$ diagonalizes the tangent vector while leaving I_n fixed. We can now use the procedure above to compute the geodesic $\tilde{\gamma}$ with initial point $\tilde{\gamma}(0) = I_n$ and tangent vector $\tilde{\gamma}'(0) = \Sigma$. Finally, the result is mapped back to the original configuration by the inverse group action, ϕ_{gv}. That is,

$$\gamma(t) = \phi_{gv}(\tilde{\gamma}(t)) = (gv)\exp(t\Sigma)(gv)^T.$$

If we flow to $t = 1$ along the geodesic γ we get the Riemannian exponential map at p (denoted Exp_p, and not to be confused with the matrix exponential map), that is,

$$\mathrm{Exp}_p(X) = \gamma(1).$$

In summary we have

Algorithm 1: Riemannian Exponential Map
Input: Initial point $p \in P(n)$.
    Tangent vector $X \in \mathrm{Sym}(n)$.
Output: $\mathrm{Exp}_p(X)$
  Let $p = u\Lambda u^T$ ($u \in SO(n)$, Λ diagonal)
  $g = u\sqrt{\Lambda}$
  $Y = g^{-1}X(g^{-1})^T$
  Let $Y = v\Sigma v^T$ ($v \in SO(n)$, Σ diagonal)
  $\mathrm{Exp}_p(X) = (gv)\exp(\Sigma)(gv)^T$

An important property of the geodesics in $P(n)$ under this metric is that they are infinitely extendible, i.e., the geodesic $\gamma(t)$ is defined for $-\infty < t < \infty$. A manifold with this property is called *complete*. Again, Fig. 1 demonstrates that the symmetric space geodesic γ remains within $P(2)$ for all t. In contrast the straight line l quickly leaves the space $P(2)$.

The map Exp_p has an inverse, called the Riemannian log map and denoted Log_p. It maps a point $x \in P(n)$ to the unique tangent vector at p that is the initial velocity of the unique geodesic γ with $\gamma(0) = p$ and $\gamma(1) = x$. Using a similar diagonalization procedure, the log map is computed by

Algorithm 2: Riemannian Log Map
Input: Initial point $p \in P(n)$.
End point $x \in P(n)$.
Output: $\mathrm{Log}_p(x)$
Let $p = u\Lambda u^T$ ($u \in SO(n)$, Λ diagonal)
$g = u\sqrt{\Lambda}$
$y = g^{-1}x(g^{-1})^T$
Let $y = v\Sigma v^T$ ($v \in SO(n)$, Σ diagonal)
$\mathrm{Log}_p(x) = (gv)\log(\Sigma)(gv)^T$

Using the notation of Algorithm 2, geodesic distance between the diffusion tensors $p, x \in P(n)$ is computed by $d(p,x) = ||\,\mathrm{Log}_p(x)||_p = \mathrm{tr}(\log(\Sigma)^2)$.

4 Statistics of Diffusion Tensors

Having formulated the geometry of diffusion tensors as a symmetric space, we now develop methods for computing statistics in this nonlinear space.

4.1 Averages of Diffusion Tensors

To define an average of diffusion tensors we follow Fréchet [9], who defines the mean of a random variable in an arbitrary metric space as the point that minimizes the expected value of the sum-of-squared distance function. Consider a set of points $A = \{x_1, \ldots, x_N\}$ on a Riemannian manifold M. Then we will be concerned with the sum-of-squared distance function

$$\rho_A(x) = \frac{1}{2N}\sum_{i=1}^{N} d(\mu, x_i)^2,$$

where d is geodesic distance on M. The *intrinsic mean* of the points in A is defined as a minimum of ρ_A, that is,

$$\mu = \arg\min_{x \in M} \rho_A(x). \tag{2}$$

The properties of the intrinsic mean have been studied by Karcher [11], and Pennec [13] describes a gradient descent algorithm to compute the mean. Since the mean is given by the minimization problem (2), we must verify that such a minimum exists and is unique. Karcher shows that for a manifold with non-positive sectional curvature the mean is uniquely defined. In fact, the space $P(n)$ does have non-positive sectional curvature, and, thus, the mean is uniquely defined. Also, the gradient of ρ_A is given by

$$\nabla\rho_A(x) = -\frac{1}{N}\sum_{i=1}^{N}\mathrm{Log}_x(x_i)$$

Thus the intrinsic mean of a collection of diffusion tensors is computed by the following gradient descent algorithm:

Algorithm 3: Intrinsic Mean of Diffusion Tensors
Input: $p_1, \ldots, p_N \in P(n)$
Output: $\mu \in P(n)$, the intrinsic mean

$\mu_0 = I$
Do
$X_i = \frac{1}{N}\sum_{k=1}^{N}\mathrm{Log}_{\mu_i}(p_k)$
$\mu_{i+1} = \mathrm{Exp}_{\mu_i}(X_i)$
While $||X_i|| > \epsilon$.

4.2 Principal Geodesic Analysis

Principal component analysis (PCA) is a useful method for describing the variability of Euclidean data. In our previous work [8] we introduced *principal geodesic analysis* (PGA) as a generalization of PCA to study the variability of data in a Lie group. In this section we review the method of principal geodesic analysis and apply it to the symmetric space of diffusion tensors. We begin with a review of PCA in Euclidean space. Consider a set of points $x_1, \ldots, x_N \in \mathbb{R}^d$ with zero mean. Principal component analysis seeks a sequence of linear subspaces that best represent the variability of the data. To be more precise, the intent is to find a orthonormal basis $\{v_1, \ldots, v_d\}$ of $\mathbb{R}^d$, which satisfies the recursive relationship

$$v_1 = \underset{||v||=1}{\arg\max}\sum_{i=1}^{N}\langle v, x_i\rangle^2, \tag{3}$$

$$v_k = \underset{||v||=1}{\arg\max}\sum_{i=1}^{N}\sum_{j=1}^{k-1}\langle v_j, x_i\rangle^2 + \langle v, x_i\rangle^2. \tag{4}$$

In other words, the subspace $V_k = \mathrm{span}(\{v_1, \ldots, v_k\})$ is the k-dimensional subspace that maximizes the variance of the data projected to that subspace. The basis $\{v_k\}$ is computed as the set of eigenvectors of the sample covariance matrix of the data.

Now turning to manifolds, consider a set of points $p_1, \ldots, p_N$ on a Riemannian manifold M. Our goal is to describe the variability of the p_i in a way that is analogous to PCA. Thus we will project the data onto lower-dimensional subspaces that best represent the variability of the data. This requires first extending three important concepts of PCA into the manifold setting:

– **Variance.** Following the work of Fréchet, we define the sample variance of the data as the expected value of the squared Riemannian distance from the mean.
– **Geodesic subspaces.** The lower-dimensional subspaces in PCA are linear subspaces. For manifolds we extend the concept of a linear subspace to that of a *geodesic submanifold*.
– **Projection.** In PCA the data is projected onto linear subspaces. We define a projection operator for geodesic submanifolds, and show how it may be efficiently approximated.

We now develop each of these concepts in detail.

Variance The variance σ^2 of a real-valued random variable x with mean μ is given by the formula $\sigma^2 = \mathcal{E}[(x-\mu)^2]$, where $\mathcal{E}$ denotes expectation. It measures the expected localization of the variable x about the mean. The definition of variance we use comes from Fréchet [9], who defines the variance of a random variable in a metric space as the expected value of the squared distance from the mean. That is, for a random variable x in a metric space with intrinsic mean μ, the variance is given by

$$\sigma^2 = \mathcal{E}[d(\mu, x)^2].$$

Thus in the manifold case, given data points $p_1, \ldots, p_N \in M$ with mean μ, we define the sample variance of the data as

$$\sigma^2 = \sum_{i=1}^{N} d(\mu, p_i)^2 = \sum_{i=1}^{N} || \operatorname{Log}_\mu(p_i)||^2. \tag{5}$$

Notice that if M is $\mathbb{R}^n$, then the variance definition in (5) is given by the trace of the sample covariance matrix, i.e., the sum of its eigenvalues. It is in this sense that this definition captures the total variation of the data.

Geodesic Submanifolds The next step in generalizing PCA to manifolds is to generalize the notion of a linear subspace. A geodesic is a curve that is locally the shortest path between points. In this way a geodesic is the generalization of a straight line. Thus it is natural to use a geodesic curve as the one-dimensional subspace, i.e., the analog of the first principal direction in PCA.

In general if N is a submanifold of a manifold M, geodesics of N are not necessarily geodesics of M. For instance the sphere S^2 is a submanifold of $\mathbb{R}^3$, but its geodesics are great circles, while geodesics of $\mathbb{R}^3$ are straight lines. A submanifold H of M is said to be *geodesic at* $x \in H$ if all geodesics of H passing through x are also geodesics of M. For example, a linear subspace of $\mathbb{R}^d$ is a submanifold geodesic at 0. Submanifolds geodesic at x preserve distances to x. This is an essential property for PGA because variance is defined by squared distance to the mean. Thus submanifolds geodesic at the mean will be the generalization of the linear subspaces of PCA.

Projection The projection of a point $x \in M$ onto a geodesic submanifold H of M is defined as the point on H that is nearest to x in Riemannian distance. Thus we define the projection operator $\pi_H : M \to H$ as

$$\pi_H(x) = \arg\min_{y \in H} d(x, y)^2.$$

Since projection is defined by a minimization, there is no guarantee that the projection of a point exists or that it is unique. However, because $P(n)$ has non-positive curvature and no conjugate points, projection onto geodesic submanifolds is unique in this case.

Projection onto a geodesic submanifold at μ can be approximated in the tangent space to the mean, $T_\mu M$. If $v_1, \ldots, v_k$ is an orthonormal basis for $T_\mu H$, then the projection operator can be approximated by the formula

$$\mathrm{Log}_\mu \left(\pi_H(x)\right) \approx \sum_{i=1}^{k} \langle v_i, \mathrm{Log}_\mu(x) \rangle_\mu. \tag{6}$$

4.3 Computing Principal Geodesic Analysis

We are now ready to define principal geodesic analysis for data $p_1, \ldots, p_N$ on a connected Riemannian manifold M. Our goal, analogous to PCA, is to find a sequence of nested geodesic submanifolds that maximize the projected variance of the data. These submanifolds are called the *principal geodesic submanifolds*.

The principal geodesic submanifolds are defined by first constructing an orthonormal basis of tangent vectors $v_1, \ldots, v_d$ that span the tangent space $T_\mu M$. These vectors are then used to form a sequence of nested subspaces $V_k = \mathrm{span}(\{v_1, \ldots, v_k\})$. The principal geodesic submanifolds are the images of the V_k under the exponential map: $H_k = \mathrm{Exp}_\mu(V_k)$. The first principal direction is chosen to maximize the projected variance along the corresponding geodesic:

$$\begin{gathered} v_1 = \arg\max_{||v||=1} \sum_{i=1}^{N} || \mathrm{Log}_\mu(\pi_H(p_i))||^2, \\ \text{where} \quad H = \exp(\mathrm{span}(\{v\})). \end{gathered} \tag{7}$$

The remaining principal directions are then defined recursively as

$$\begin{gathered} v_k = \arg\max_{||v||=1} \sum_{i=1}^{N} || \mathrm{Log}_\mu(\pi_H(p_i))||^2, \\ \text{where} \quad H = \exp(\mathrm{span}(\{v_1, \ldots, v_{k-1}, v\})). \end{gathered} \tag{8}$$

If we use (6) to approximate the projection operator π_H in (7) and (8), we get

$$v_1 \approx \arg\max_{||v||=1} \sum_{i=1}^{N} \langle v, \mathrm{Log}_\mu(p_i) \rangle_\mu^2,$$

$$v_k \approx \arg\max_{||v||=1} \sum_{i=1}^{N} \sum_{j=1}^{k-1} \langle v_j, \mathrm{Log}_\mu(p_i) \rangle_\mu^2 + \langle v, \mathrm{Log}_\mu(p_i) \rangle_\mu^2.$$

The above minimization problem is simply the standard principal component analysis in $T_\mu M$ of the vectors $\mathrm{Log}_\mu(p_i)$, which can be seen by comparing the approximations above to the PCA equations, (3) and (4). Applying these ideas to $P(n)$, we have the following algorithm for approximating the PGA of diffusion tensor data:

Algorithm 4: PGA of Diffusion Tensors
Input: $p_1, \ldots, p_N \in P(n)$
Output: Principal directions, $v_k \in \mathrm{Sym}(n)$
Variances, $\lambda_k \in \mathbb{R}$
$\mu =$ intrinsic mean of $\{p_i\}$ (Algorithm 3)
$x_i = \mathrm{Log}_\mu(p_i)$
$\mathbf{S} = \frac{1}{N}\sum_{i=1}^{N} x_i x_i^T$ (treating the x_i as column vectors)
$\{v_k, \lambda_k\}$ = eigenvectors/eigenvalues of $\mathbf{S}$.

A new diffusion tensor p can now be generated from the PGA by the formula $p = \mathrm{Exp}_\mu\left(\sum_{k=1}^{d} \alpha_k v_k\right)$, where the $\alpha_k \in \mathbb{R}$ are the coefficients of the modes of variation.

5 Properties of PGA on $P(n)$

We now demonstrate that PGA on the symmetric space $P(n)$ preserves certain important properties of the diffusion tensor data, namely the properties of positive-definiteness, determinant, and orientation. This makes the symmetric space formulation an attractive approach for the statistical analysis of diffusion tensor images. We have already mentioned that, in contrast to linear PCA, symmetric space PGA preserves positive-definiteness. That is, the principal geodesics are completely contained within $P(n)$, and any matrix generated by the principal geodesics will be positive-definite.

The next two properties we consider are the determinant and orientation. Consider a collection of diffusion tensors that all have the same determinant D. We wish to show that the resulting average and any tensor generated by the principal geodesic analysis will also have determinant D. To show this we first look at the subset of $P(n)$ of matrices with determinant D, that is, the subset $P_D = \{p \in P(n) : \det(p) = D\}$. This subset is a *totally geodesic submanifold*, meaning that any geodesic within P_D is a geodesic of the full space $P(n)$. Notice the difference from the definition of a submanifold geodesic at a point; totally geodesic submanifolds are geodesic at *every* point. Now, the fact that P_D is totally geodesic implies that the averaging process in Algorithm 3 will remain in P_D if all the data lies in P_D. Also, the principal directions v_k in the PGA will lie in the tangent subspace $T_\mu P_D$. Thus any diffusion tensor generated by the principal geodesics will remain in the space P_D.

The same argument may be applied to show that symmetric space averaging and PGA preserve the orientation of diffusion tensors. In fact, the subset of all diffusion tensors having the same orientation is also a totally geodesic submanifold, and the same reasoning applies. Unlike the positive-definiteness and determinant, orientations are also preserved by linear averaging and PCA.

To demonstrate these properties, we simulated random 3D diffusion tensors and computed both their linear and symmetric space statistics. We first tested the determinant preservation by generating 100 random 3D diffusion tensors with determinant 1.

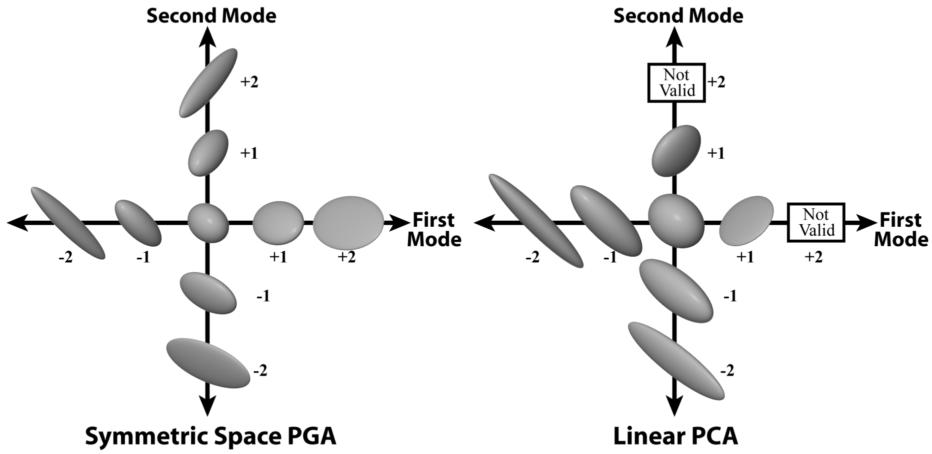

Fig. 2. The first two modes of variation of the simulated data: (left) using the symmetric space PGA, and (right) using linear PCA. Units are in standard deviations. The boxes labelled "Not Valid" indicate that the tensor was not positive-definite, i.e., it had negative eigenvalues.

To do this we first generated 100 random 3×3 symmetric matrices, with entries distributed according to a normal distribution, $N(0, \frac{1}{2})$. Then, we took the matrix exponential of these random symmetric matrices, thus making them positive-definite diffusion tensors. Finally, we normalized the random diffusion tensors to have determinant 1 by dividing each tensor by the cube root of its determinant. We then computed the linear average and PCA and symmetric space average and PGA of the simulated tensors. The results are shown in Fig. 2 as the diffusion tensors generated by the first two modes of variation. The linear PCA generated invalid diffusion tensors, i.e., tensors with negative eigenvalues, at $+2$ standard deviations in both the first and second modes. All of the diffusion tensors generated by the symmetric space PGA have determinant 1. The linear mean demonstrates the "swelling" effect of linear averaging. It has determinant 2.70, and the linear PCA tensors within ± 2 standard deviations have determinants ranging from -2.80 to 2.82. The negative determinants came from the tensors that were not positive-definite. Therefore, we see that the symmetric space PGA has preserved the positive-definiteness and the determinant, while the linear PCA has preserved neither.

Next we tested the orientation preservation by generating 100 random, axis-aligned, 3D diffusion tensors. This was done by generating 3 random eigenvalues for each matrix, corresponding to the x, y, and z axes. The eigenvalues were chosen from a log-normal distribution with log mean 0 and log standard deviation 0.5. Next we generated a random orientation $u \in SO(3)$ and applied it to all of the axis-aligned matrices by the map $p \mapsto upu^T$. Thus each of the diffusion tensors in our test set had eigenvectors equal to the columns of the rotation matrix u. We computed both the symmetric space and linear statistics of the data. As was expected, both methods preserved the orientations. However, the linear PCA again generated tensors that were not positive-definite.

6 Conclusion

We have presented a framework for the statistical analysis of diffusion tensor images. The methods rely on regarding the space of diffusion tensors as a Riemannian symmetric space. We developed algorithms for computing averages and modes of variation of diffusion tensor data by extending statistical methods to the symmetric space setting. The methods presented in this paper lay the groundwork for statistical studies of the variability of diffusion tensor images across patients.

7 Acknowledgements

We would like to thank Dr. Guido Gerig and Dr. Stephen Pizer for useful discussions and suggestions. This work was done with support from NIH grants P01 CA47982 and EB02779.

References

[1] D. C. Alexander, C. Pierpaoli, P.J. Basser, and J.C. Gee. Spatial transformations of diffusion tensor MR images. *IEEE Transactions on Medical Imaging*, 20(11):1131–1139, 2001.

[2] P. J. Basser, J. Mattiello, and D. Le Bihan. MR diffusion tensor spectroscopy and imaging. *Biophysics Journal*, 66:259–267, 1994.

[3] P. J. Basser and S. Pajevic. A normal distribution for tensor-valued random variables: applications to diffusion tensor MRI. *IEEE Transactions on Medical Imaging*, 22(7):785–794, 2003.

[4] D. Le Bihan, J.-F. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko, and H. Chabriat. Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging*, 13:534–546, 2001.

[5] W. M. Boothby. *An Introduction to Differentiable Manifolds and Riemannian Geometry*. Academic Press, 2nd edition, 1986.

[6] C. Chefd'hotel, D. Tschumperlé, R. Deriche, and O. Faugeras. Constrained flows of matrix-valued functions: Application to diffusion tensor regularization. In *European Conference on Computer Vision*, pages 251–265, 2002.

[7] O. Coulon, D. C. Alexander, and S. Arridge. Diffusion tensor magnetic resonance image regularization. *Medical Image Analysis*, 8(1):47–68, 2004.

[8] P. T. Fletcher, C. Lu, and S. Joshi. Statistics of shape via principal geodesic analysis on Lie groups. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 95–101, 2003.

[9] M. Fréchet. Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. Inst. H. Poincaré*, (10):215–310, 1948.

[10] S. Helgason. *Differential Geometry, Lie Groups, and Symmetric Spaces*. Academic Press, 1978.

[11] H. Karcher. Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics*, 30(5):509–541, 1977.

[12] S. Pajevic and P. J. Basser. Parametric and non-parametric statistical analysis of DT-MRI. *Journal of Magnetic Resonance*, 161(1):1–14, 2003.

[13] X. Pennec. Probabilities and statistics on Riemannian manifolds: basic tools for geometric measurements. In *IEEE Workshop on Nonlinear Signal and Image Processing*, 1999.

[14] Z. Wang, B. C. Vemuri, Y. Chen, and T. Mareci. A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from DWI. In *Information Processing in Medical Imaging*, pages 660–671, 2003.