Statistics on Riemannian Manifolds

Tom Fletcher
Scientific Computing and Imaging Institute
University of Utah

August 19, 2009

THE BRAIN INSTITUTE
THE UNIVERSITY OF UTAH

SCI INSTITUTE

Manifold Data

“Learned” Manifolds

“Known” Manifolds

Manifold Data

“Learned” Manifolds

- Raw data lies in Euclidean space
- Manifold + Noise

“Known” Manifolds

Manifold Data

“Learned” Manifolds

- Raw data lies in Euclidean space
- Manifold + Noise

“Known” Manifolds

- Raw data lies in a manifold
- Typically given by some constraints on data

Manifold Data in Vision and Imaging

- Directional data

Manifold Data in Vision and Imaging

- Directional data
- Transformation groups (rotations, projective, affine)

Manifold Data in Vision and Imaging

- Directional data
- Transformation groups (rotations, projective, affine)
- Shapes

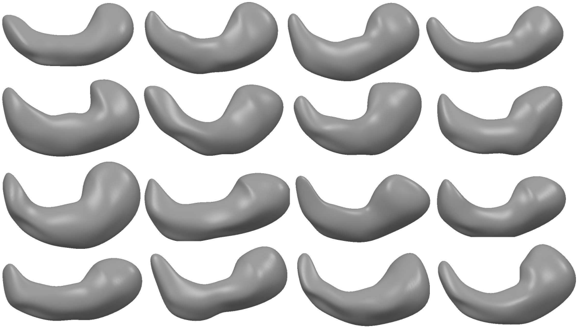
Manifold Data in Vision and Imaging

- Directional data
- Transformation groups (rotations, projective, affine)
- Shapes
- Diffusion tensors, structure tensors

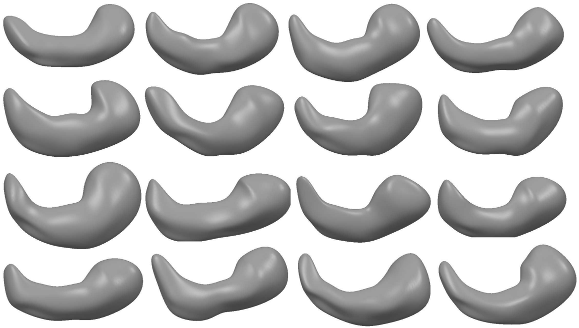
Manifold Data in Vision and Imaging

- Directional data
- Transformation groups (rotations, projective, affine)
- Shapes
- Diffusion tensors, structure tensors
- Diffeomorphisms (for deformable atlas building)

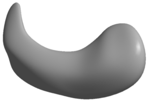
Manifold Statistics: Averages



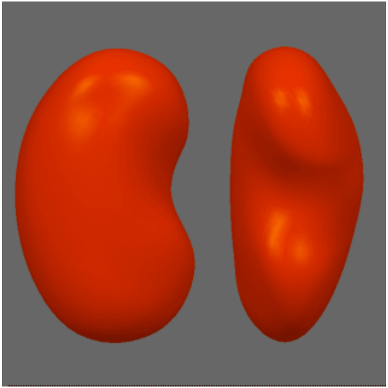
Manifold Statistics: Averages

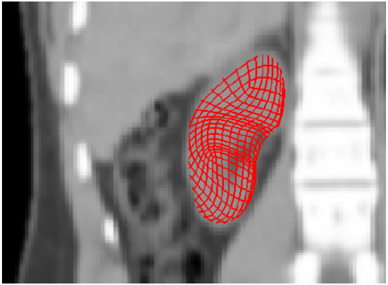


$\longrightarrow$



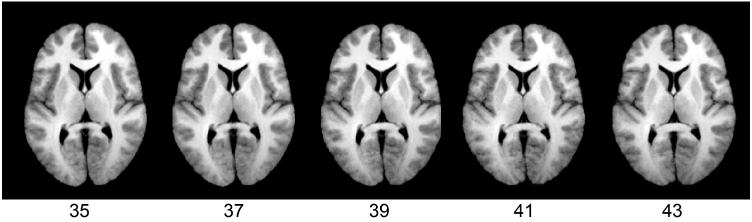
Manifold Statistics: Variability

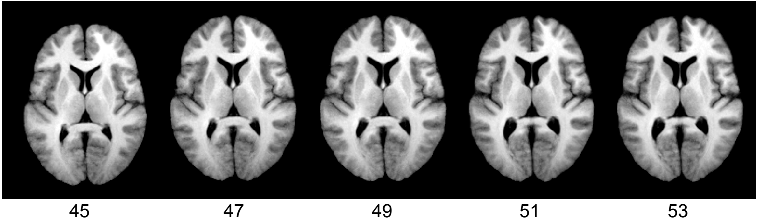




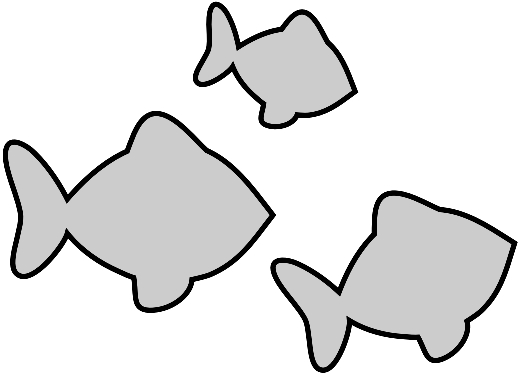
Shape priors in segmentation

Manifold Statistics: Regression

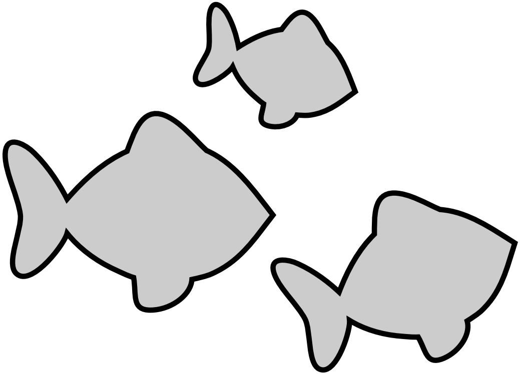




What is Shape?



What is Shape?

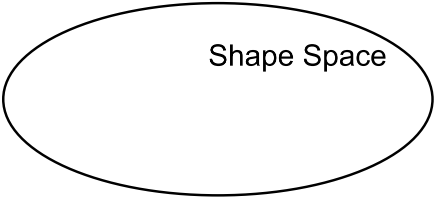


Shape is the geometry of an object modulo position, orientation, and size.

Shape Representations

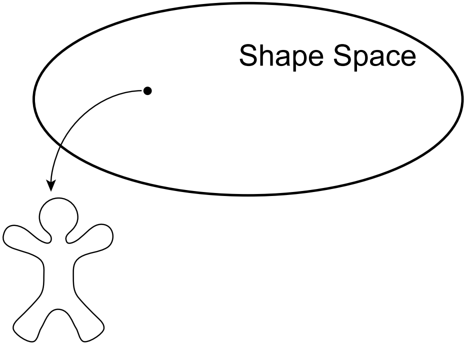
- Boundary models (points, curves, surfaces, level sets)
- Interior models (medial, solid mesh)
- Transformation models (splines, diffeomorphisms)

Shape Analysis



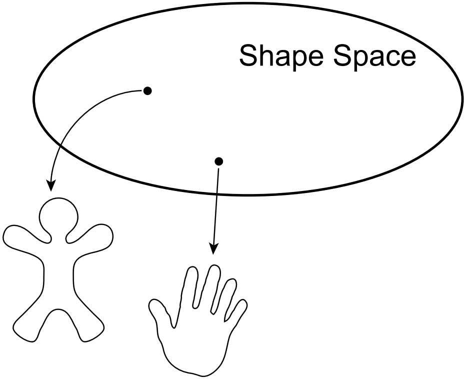
A shape is a point in a high-dimensional, nonlinear shape space.

Shape Analysis



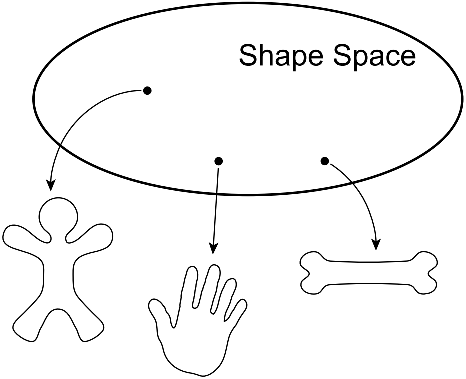
A shape is a point in a high-dimensional, nonlinear shape space.

Shape Analysis



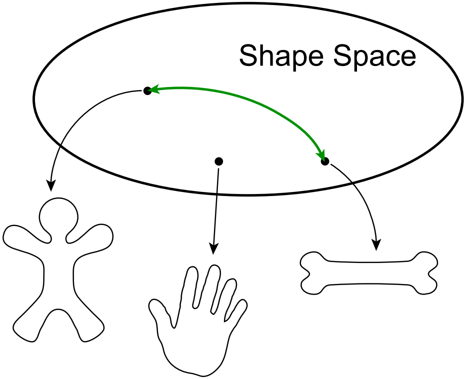
A shape is a point in a high-dimensional, nonlinear shape space.

Shape Analysis



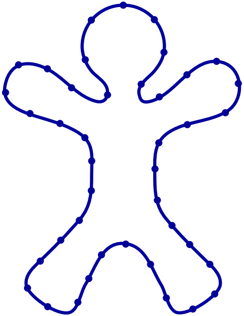
A shape is a point in a high-dimensional, nonlinear shape space.

Shape Analysis



A metric space structure provides a comparison between two shapes.

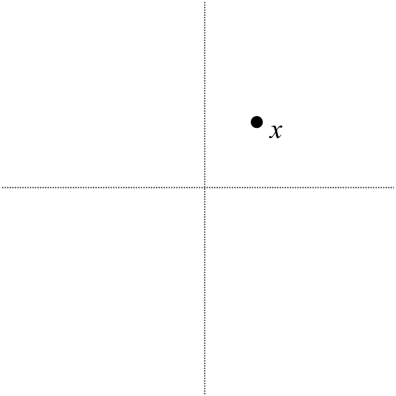
Kendall's Shape Space



- Define object with k points.
- Represent as a vector in $\mathbb{R}^{2k}$.
- Remove translation, rotation, and scale.
- End up with complex projective space, $\mathbb{CP}^{k-2}$.

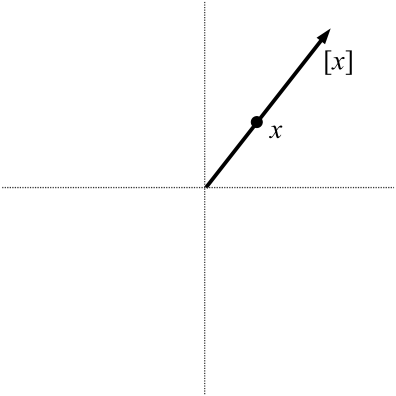
Quotient Spaces

What do we get when we “remove” scaling from $\mathbb{R}^2$?



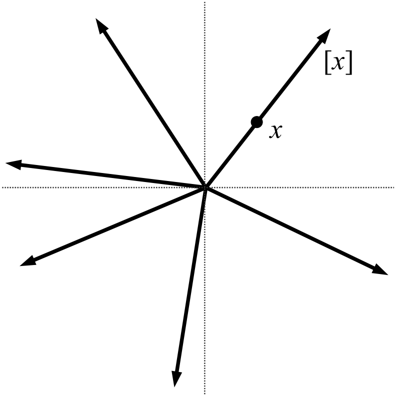
Quotient Spaces

What do we get when we "remove" scaling from $\mathbb{R}^2$?



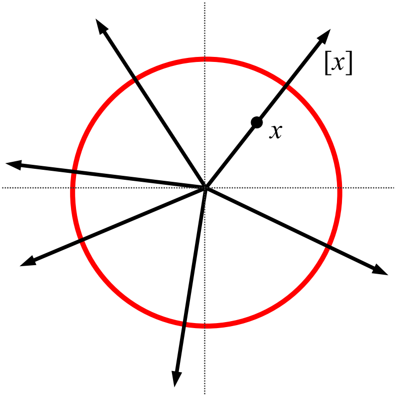
Quotient Spaces

What do we get when we "remove" scaling from $\mathbb{R}^2$?



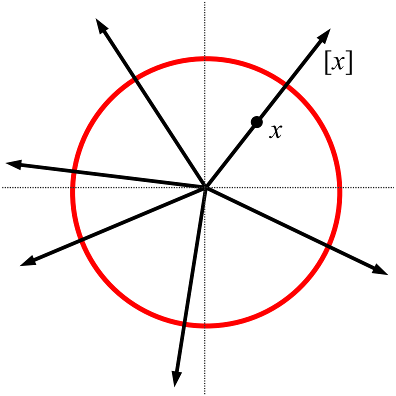
Quotient Spaces

What do we get when we “remove” scaling from $\mathbb{R}^2$?



Quotient Spaces

What do we get when we “remove” scaling from $\mathbb{R}^2$?



Notation: $[x] \in \mathbb{R}^2/\mathbb{R}^+$

Constructing Kendall's Shape Space

- Consider planar landmarks to be points in the complex plane.

Constructing Kendall's Shape Space

- Consider planar landmarks to be points in the complex plane.
- An object is then a point $(z_1, z_2, \ldots, z_k) \in \mathbb{C}^k$.

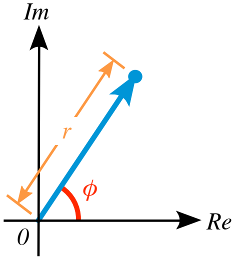
Constructing Kendall's Shape Space

- Consider planar landmarks to be points in the complex plane.
- An object is then a point $(z_1, z_2, \ldots, z_k) \in \mathbb{C}^k$.
- Removing **translation** leaves us with $\mathbb{C}^{k-1}$.

Constructing Kendall's Shape Space

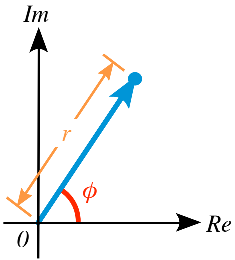
- Consider planar landmarks to be points in the complex plane.
- An object is then a point $(z_1, z_2, \ldots, z_k) \in \mathbb{C}^k$.
- Removing **translation** leaves us with $\mathbb{C}^{k-1}$.
- How to remove **scaling** and **rotation**?

Scaling and Rotation in the Complex Plane



Recall a complex number can be written as $z = re^{i\phi}$, with modulus r and argument ϕ.

Scaling and Rotation in the Complex Plane



Recall a complex number can be written as $z = re^{i\phi}$, with modulus r and argument ϕ.

Complex Multiplication:

$$se^{i\theta} * re^{i\phi} = (sr)e^{i(\theta+\phi)}$$

Multiplication by a complex number $se^{i\theta}$ is equivalent to scaling by s and rotation by θ.

Removing Scale and Translation

Multiplying a centered point set, $\mathbf{z} = (z_1, z_2, \ldots, z_{k-1})$, by a constant $w \in \mathbb{C}$, just rotates and scales it.

Removing Scale and Translation

Multiplying a centered point set, $\mathbf{z} = (z_1, z_2, \ldots, z_{k-1})$, by a constant $w \in \mathbb{C}$, just rotates and scales it.

Thus the shape of $\mathbf{z}$ is an equivalence class:

$$[\mathbf{z}] = \{(wz_1, wz_2, \ldots, wz_{k-1}) : \forall w \in \mathbb{C}\}$$

Removing Scale and Translation

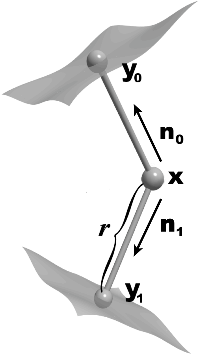
Multiplying a centered point set, $\mathbf{z} = (z_1, z_2, \ldots, z_{k-1})$, by a constant $w \in \mathbb{C}$, just rotates and scales it.

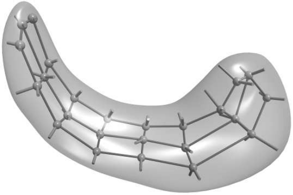
Thus the shape of $\mathbf{z}$ is an equivalence class:

$$[\mathbf{z}] = \{(wz_1, wz_2, \ldots, wz_{k-1}) : \forall w \in \mathbb{C}\}$$

This gives complex projective space $\mathbb{CP}^{k-2}$ – much like the sphere comes from equivalence classes of scalar multiplication in $\mathbb{R}^n$.

The M-rep Shape Space





Medial Atom:

$$\mathbf{m} = \{\mathbf{x}, r, \mathbf{n}_0, \mathbf{n}_1\} \in \mathcal{M}(1)$$

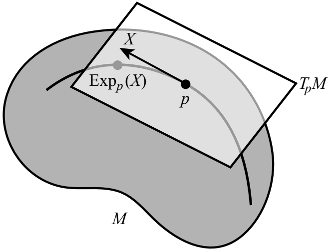
$$\mathcal{M}(1) = \mathbb{R}^3 \times \mathbb{R}^+ \times S^2 \times S^2$$

M-rep Model with n atoms:

$$\mathbf{M} \in \mathcal{M}(n) = \mathcal{M}(1)^n$$

Shape change in terms of local translation, bending, & widening.

The Exponential and Log Maps



- The exponential map takes tangent vectors to points along geodesics.
- The length of the tangent vector equals the length along the geodesic segment.
- Its inverse is the log map – it gives distance between points: $d(p, q) = \| \mathrm{Log}_p(q) \|$.

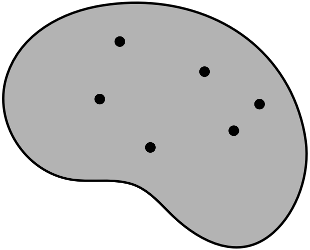
Intrinsic Means (Fréchet)

The *intrinsic mean* of a collection of points $x_1, \ldots, x_N$ on a Riemannian manifold M is

$$\mu = \arg\min_{x \in M} \sum_{i=1}^{N} d(x, x_i)^2,$$

where $d(\cdot, \cdot)$ denotes Riemannian distance on M.

Computing Means



Gradient Descent Algorithm:

Input: $\mathbf{x}_1, \ldots, \mathbf{x}_N \in M$

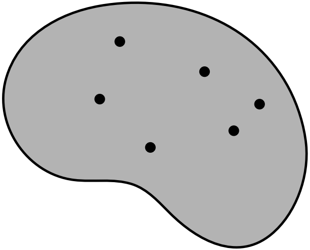
$\mu_0 = \mathbf{x}_1$

Repeat:

$$\Delta\mu = \frac{1}{N}\sum_{i=1}^{N} \mathrm{Log}_{\mu_k}(\mathbf{x}_i)$$

$$\mu_{k+1} = \mathrm{Exp}_{\mu_k}(\Delta\mu)$$

Computing Means



Gradient Descent Algorithm:

Input: $\mathbf{x}_1, \ldots, \mathbf{x}_N \in M$

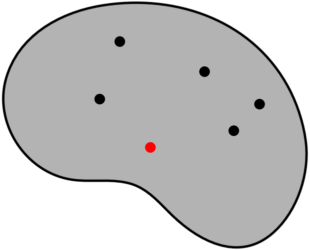
$\mu_0 = \mathbf{x}_1$

Repeat:

$$\Delta\mu = \frac{1}{N} \sum_{i=1}^{N} \mathrm{Log}_{\mu_k}(\mathbf{x}_i)$$

$$\mu_{k+1} = \mathrm{Exp}_{\mu_k}(\Delta\mu)$$

Computing Means



Gradient Descent Algorithm:

Input: $\mathbf{x}_1, \ldots, \mathbf{x}_N \in M$

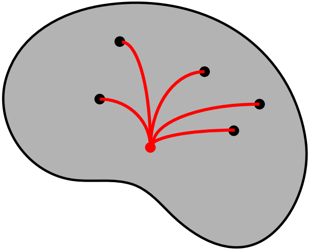
$\mu_0 = \mathbf{x}_1$

Repeat:

$$\Delta\mu = \frac{1}{N} \sum_{i=1}^{N} \mathrm{Log}_{\mu_k}(\mathbf{x}_i)$$

$$\mu_{k+1} = \mathrm{Exp}_{\mu_k}(\Delta\mu)$$

Computing Means



Gradient Descent Algorithm:

Input: $\mathbf{x}_1, \ldots, \mathbf{x}_N \in M$

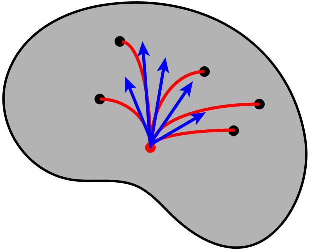
$\mu_0 = \mathbf{x}_1$

Repeat:

$$\Delta\mu = \frac{1}{N} \sum_{i=1}^{N} \mathrm{Log}_{\mu_k}(\mathbf{x}_i)$$

$$\mu_{k+1} = \mathrm{Exp}_{\mu_k}(\Delta\mu)$$

Computing Means



Gradient Descent Algorithm:

Input: $\mathbf{x}_1, \ldots, \mathbf{x}_N \in M$

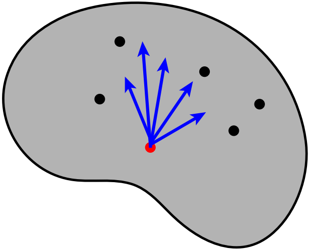
$\mu_0 = \mathbf{x}_1$

Repeat:

$$\Delta\mu = \frac{1}{N}\sum_{i=1}^{N} \mathrm{Log}_{\mu_k}(\mathbf{x}_i)$$

$$\mu_{k+1} = \mathrm{Exp}_{\mu_k}(\Delta\mu)$$

Computing Means



Gradient Descent Algorithm:

Input: $\mathbf{x}_1, \ldots, \mathbf{x}_N \in M$

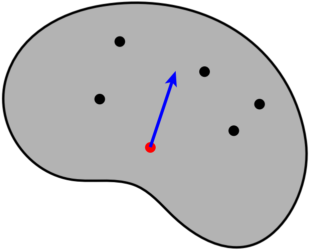
$\mu_0 = \mathbf{x}_1$

Repeat:

$$\Delta\mu = \frac{1}{N}\sum_{i=1}^{N} \mathrm{Log}_{\mu_k}(\mathbf{x}_i)$$

$$\mu_{k+1} = \mathrm{Exp}_{\mu_k}(\Delta\mu)$$

Computing Means



Gradient Descent Algorithm:

Input: $\mathbf{x}_1, \ldots, \mathbf{x}_N \in M$

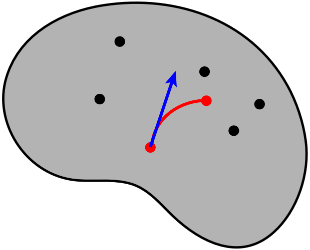
$\mu_0 = \mathbf{x}_1$

Repeat:

$$\Delta\mu = \frac{1}{N}\sum_{i=1}^{N} \mathrm{Log}_{\mu_k}(\mathbf{x}_i)$$

$$\mu_{k+1} = \mathrm{Exp}_{\mu_k}(\Delta\mu)$$

Computing Means



Gradient Descent Algorithm:

Input: $\mathbf{x}_1, \ldots, \mathbf{x}_N \in M$

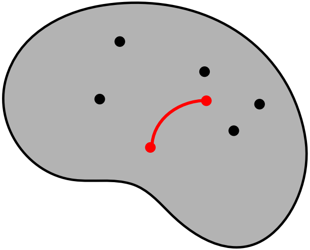
$\mu_0 = \mathbf{x}_1$

Repeat:

$$\Delta\mu = \frac{1}{N} \sum_{i=1}^{N} \mathrm{Log}_{\mu_k}(\mathbf{x}_i)$$

$$\mu_{k+1} = \mathrm{Exp}_{\mu_k}(\Delta\mu)$$

Computing Means



Gradient Descent Algorithm:

Input: $\mathbf{x}_1, \ldots, \mathbf{x}_N \in M$

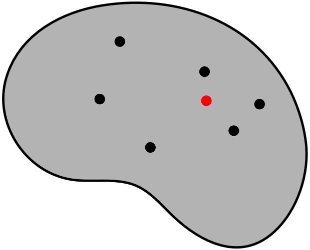
$\mu_0 = \mathbf{x}_1$

Repeat:

$$\Delta\mu = \frac{1}{N} \sum_{i=1}^{N} \mathrm{Log}_{\mu_k}(\mathbf{x}_i)$$

$$\mu_{k+1} = \mathrm{Exp}_{\mu_k}(\Delta\mu)$$

Computing Means



Gradient Descent Algorithm:

Input: $\mathbf{x}_1, \ldots, \mathbf{x}_N \in M$

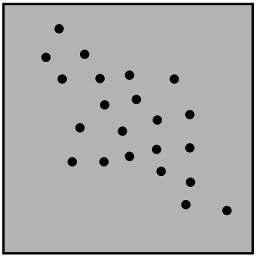
$\mu_0 = \mathbf{x}_1$

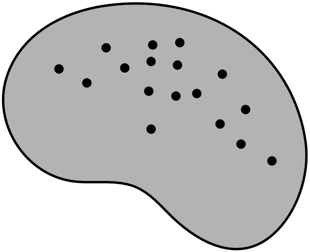
Repeat:

$$\Delta\mu = \frac{1}{N}\sum_{i=1}^{N} \mathrm{Log}_{\mu_k}(\mathbf{x}_i)$$

$$\mu_{k+1} = \mathrm{Exp}_{\mu_k}(\Delta\mu)$$

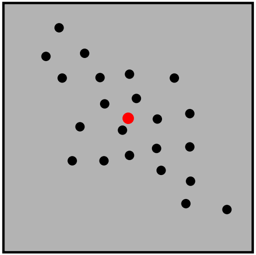
Principal Geodesic Analysis



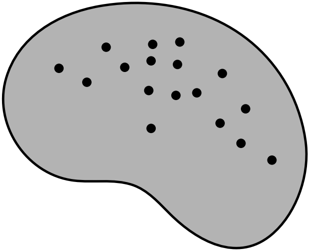


Principal Geodesic Analysis

Linear Statistics (PCA)

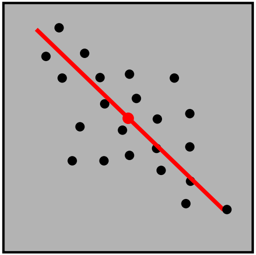


Curved Statistics (PGA)

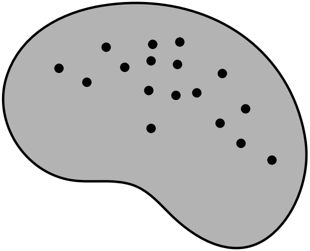


Principal Geodesic Analysis

Linear Statistics (PCA)

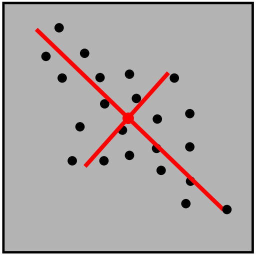


Curved Statistics (PGA)

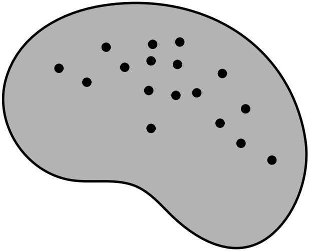


Principal Geodesic Analysis

Linear Statistics (PCA)

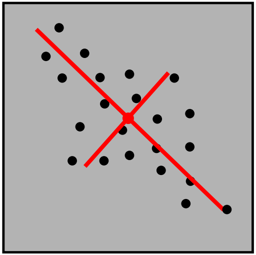


Curved Statistics (PGA)

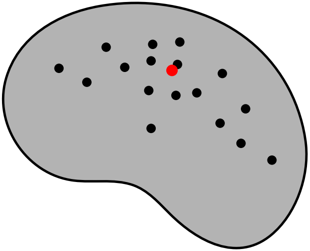


Principal Geodesic Analysis

Linear Statistics (PCA)

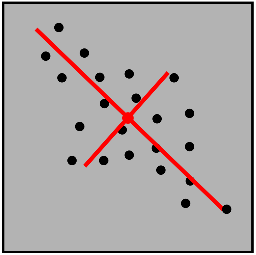


Curved Statistics (PGA)

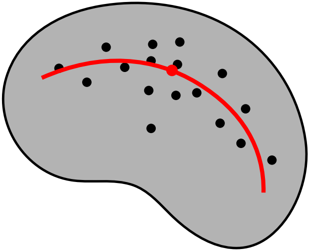


Principal Geodesic Analysis

Linear Statistics (PCA)

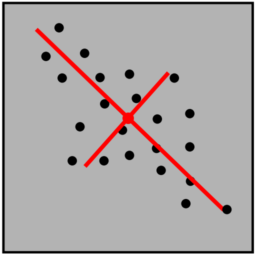


Curved Statistics (PGA)

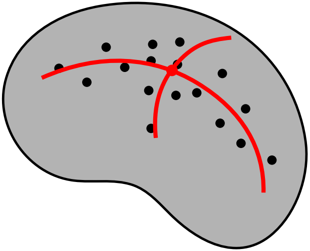


Principal Geodesic Analysis

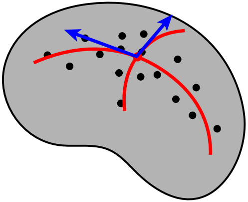
Linear Statistics (PCA)



Curved Statistics (PGA)



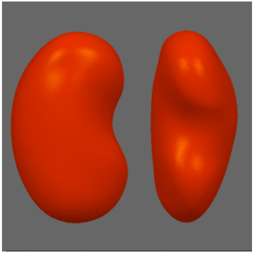
Computing PGA



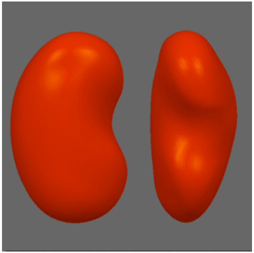
- Find nested linear subspaces $V_k \subset T_pM$ such that $\mathrm{Exp}_\mu(V_k)$ maximizes variance of projected data.
- First-order approximation: PCA in tangent space of sample covariance matrix,

$$S = \frac{1}{N-1} \sum_{i=1}^{N} \mathrm{Log}_\mu(x_i)\, \mathrm{Log}_\mu(x_i)^T$$

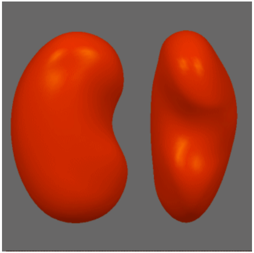
PGA of Kidney



Mode 1



Mode 2



Mode 3

Robust Statistics: Motivation

- The mean is overly influenced by outliers due to sum-of-squares.
- Robust statistical description of shape or other manifold data.
- Deal with outliers due to imaging noise or data corruption.
- Misdiagnosis, segmentation error, or outlier in a population study.

Mean vs. Median in $\mathbb{R}^n$

Mean: least-squares problem

$$\mu = \arg\min_{x \in \mathbb{R}^n} \sum \|x - x_i\|^2$$

Closed-form solution (arithmetic average)

Mean vs. Median in $\mathbb{R}^n$

Mean: least-squares problem

$$\mu = \arg\min_{x \in \mathbb{R}^n} \sum \|x - x_i\|^2$$

Closed-form solution (arithmetic average)

Geometric Median, or Fermat-Weber Point:

$$m = \arg\min_{x \in \mathbb{R}^n} \sum \|x - x_i\|$$

No closed-form solution

Weiszfeld Algorithm in $\mathbb{R}^n$

- Gradient descent on sum-of-distance:

$$m_{k+1} = m_k - \alpha G_k,$$
$$G_k = \sum_{i \in I_k} \frac{m_k - x_i}{\|x_i - m_k\|} \Bigg/ \left(\sum_{i \in I_k} \|x_i - m_k\|^{-1} \right)$$

- Step size: $0 < \alpha \leq 2$
- Exclude singular points: $I_k = \{i : m_k \neq x_i\}$
- Weiszfeld (1937), Ostresh (1978)

Geometric Median on a Manifold

The geometric median of data $x_i \in M$ is the point that minimizes the sum of geodesic distances:

$$m = \arg\min_{x \in M} \sum_{i=1}^{N} d(x, x_i)$$

Fletcher, et al. CVPR 2008 and NeuroImage 2009.

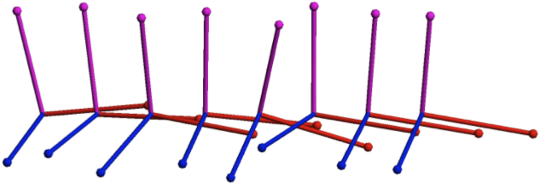
Weiszfeld Algorithm for Manifolds

Gradient descent:

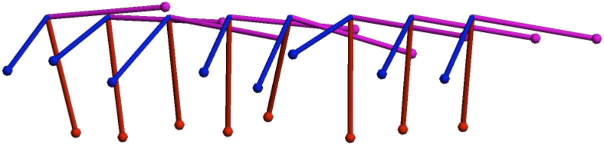
$$
\begin{aligned}
m_{k+1} &= \mathrm{Exp}_{m_k}(\alpha v_k), \\
v_k &= \sum_{i \in I_k} \frac{\mathrm{Log}_{m_k}(x_i)}{d(m_k, x_i)} \Bigg/ \left(\sum_{i \in I_k} d(m_k, x_i)^{-1} \right)
\end{aligned}
$$

Example: Rotations

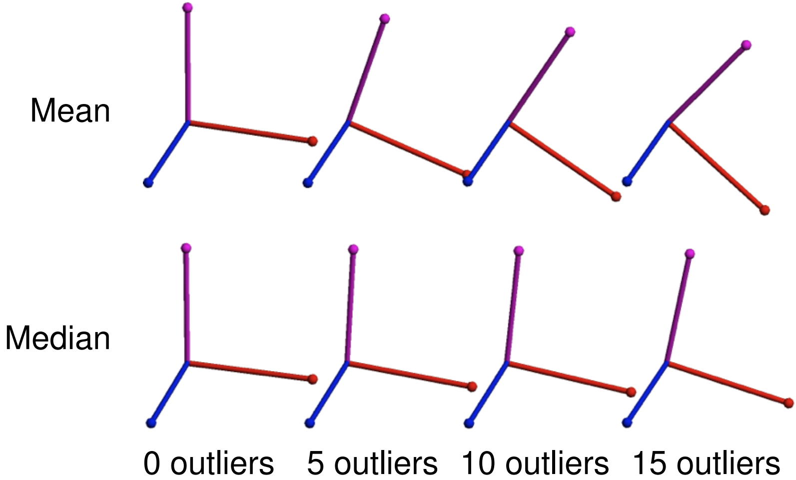
Input data: 20 random rotations



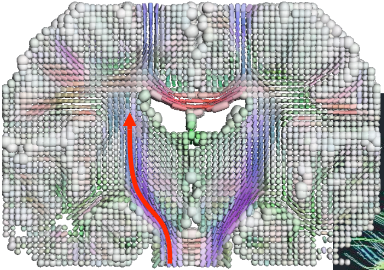
Outlier set: random, rotated 90°

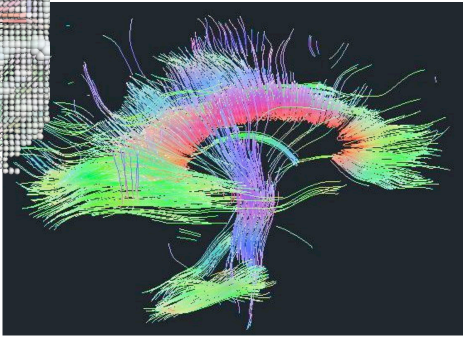


Example: Rotations



Application: Diffusion Tensor MRI





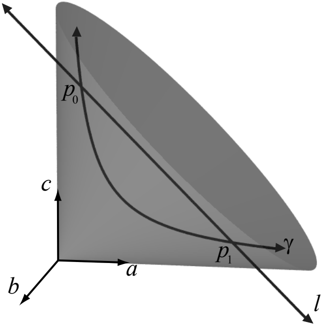
Space of Positive-Definite Tensors

- Positive-definite, symmetric matrices

$$\mathrm{PD}(n) = \mathrm{GL}^+(n)/\mathrm{SO}(n)$$

- Riemannian manifold with nonpositive curvature
- Applications:
  - Diffusion tensor MRI: Fletcher (2004), Pennec (2004)
  - Structure tensor: Rathi (2007)
  - Bookstein's simplex shape space (1986)

Example: $\mathrm{PD}(2)$



$A \in \mathrm{PD}(2)$ is of the form

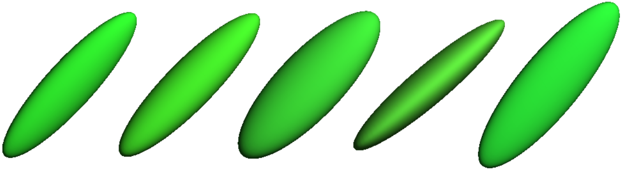
$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

$$ac - b^2 > 0, \quad a > 0.$$

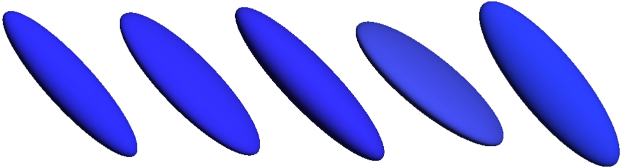
Similar situation for $\mathrm{PD}(3)$ (6-dimensional).

Example: Tensors

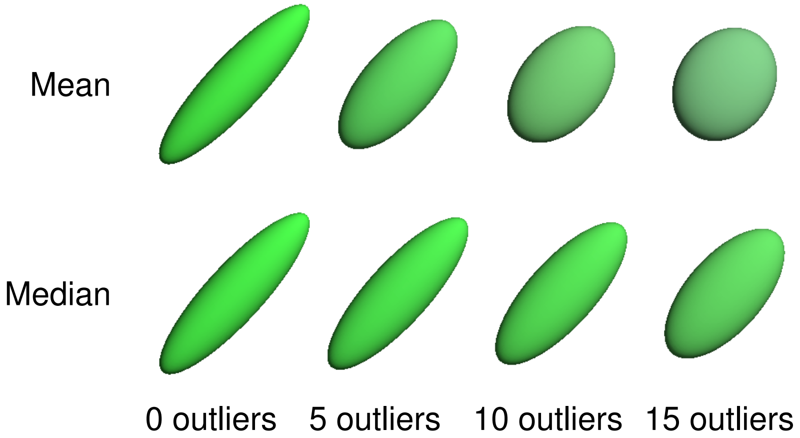
Input data: 20 random tensors



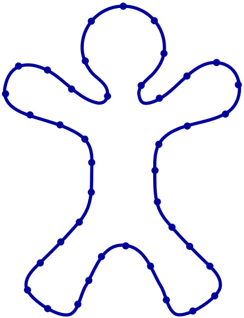
Outlier set: random, rotated 90°



Example: Tensors

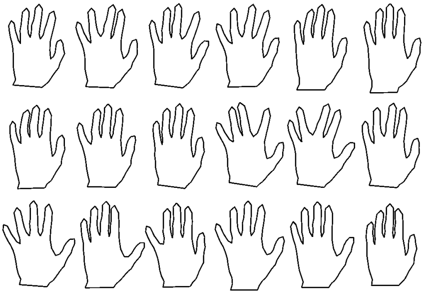


Kendall's Shape Space

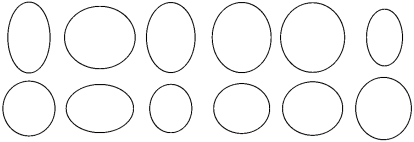


- Define object with k points.
- Represent as a vector in $\mathbb{R}^{2k}$.
- Remove translation, rotation, and scale.
- End up with complex projective space, $\mathbb{CP}^{k-2}$.

Example on Kendall Shape Spaces



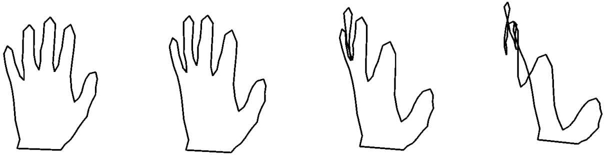
Hand shapes



Outliers

Example on Kendall Shape Spaces

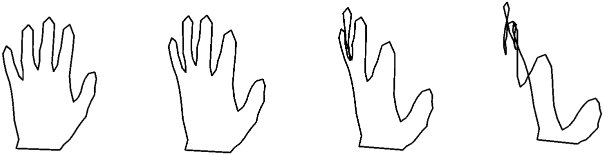
Mean:



Outliers: 0 2 6 12

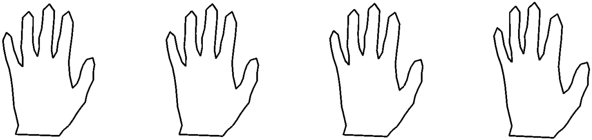
Example on Kendall Shape Spaces

Mean:



Outliers: 0 2 6 12

Median:



Outliers: 0 2 6 12

Image Metamorphosis

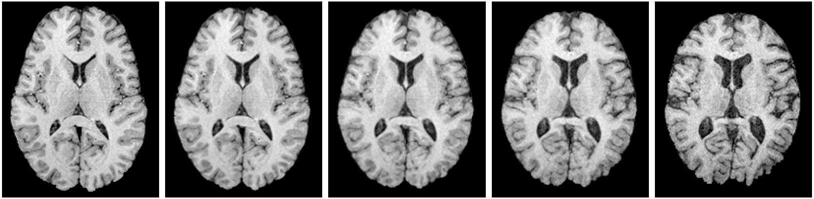
- Metric between images
- Includes both deformation and intensity change

$$U(v_t, I_t) = \int_0^1 \|v_t\|_V^2 \, dt \;+\; \frac{1}{\sigma^2} \int_0^1 \left\| \frac{dI_t}{dt} + \langle \nabla I_t, v_t \rangle \right\|_{L^2}^2 dt$$

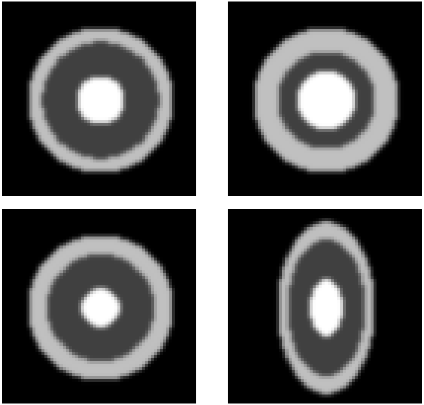
Image Metamorphosis

- Metric between images
- Includes both deformation and intensity change

$$U(v_t, I_t) = \int_0^1 \|v_t\|_V^2 \, dt \; + \; \frac{1}{\sigma^2} \int_0^1 \left\| \frac{dI_t}{dt} + \langle \nabla I_t, v_t \rangle \right\|_{L^2}^2 dt$$

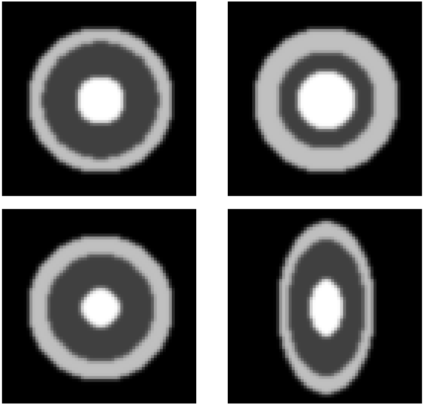


Example: Metamorphosis

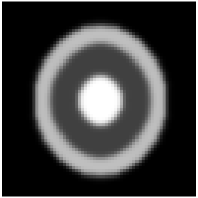


Input Data

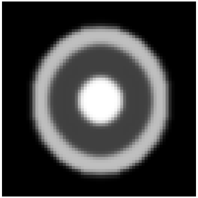
Example: Metamorphosis



Input Data

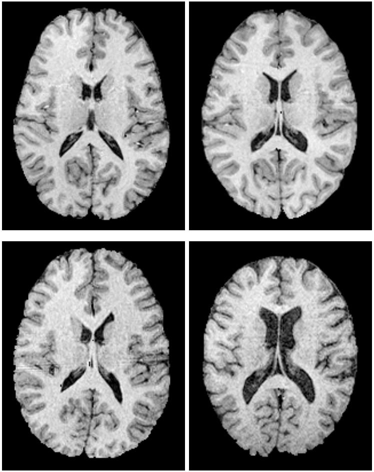


Mean
ratio = 1.13

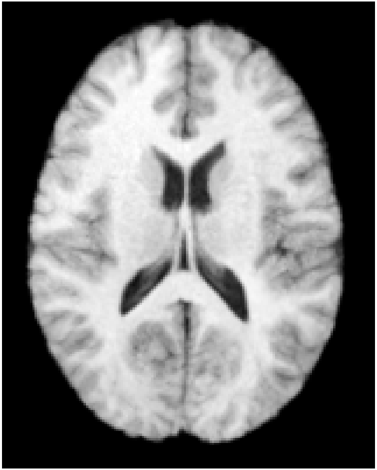


Median
ratio = 1.04

Example: Metamorphosis



Input Data



Median Atlas

Preliminaries

- $x_i \in U \subset M$, U is a convex subset
- $\mathrm{diam}(U) = \max_{x,y \in U} d(x, y)$

Existence and Uniqueness

Theorem. *The weighted geometric median exists and is unique if*

1. *the sectional curvatures of* M *are bounded above by* $\Delta > 0$ *and* $\mathrm{diam}(U) < \pi/(2\sqrt{\Delta})$, *or*
2. *the sectional curvatures of* M *are nonpositive.*

Existence and Uniqueness

Theorem. *The weighted geometric median exists and is unique if*

1. *the sectional curvatures of M are bounded above by $\Delta > 0$ and $\mathrm{diam}(U) < \pi/(2\sqrt{\Delta})$, or*
2. *the sectional curvatures of M are nonpositive.*

Proof is by showing the convexity of geodesic distance.

Identical conditions to ensure the mean (Karcher).

Robustness

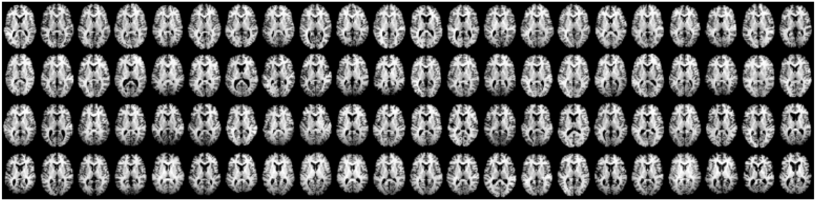
- Breakdown point: percentage of points that can be moved to infinity before statistic goes to infinity
- Euclidean mean: 0%
- Euclidean geometric median: 50%
- Same result holds for noncompact manifolds
- Does not make sense for compact manifolds

Convergence Theorem for Manifold Weiszfeld Algorithm

Theorem. *If the sectional curvatures of* M *are nonnegative and the existence/uniqueness conditions are satisfied, then* $\lim_{k\to\infty} m_k = m$ *for* $0 < \alpha \leq 2$*.*

Describing Shape Change

- How does shape change over time?
- Changes due to growth, aging, disease, etc.
- Example: 100 healthy subjects, 20–80 yrs. old



- We need regression of shape!

Regression Analysis

- Describe relationship between a dependent random variable Y to an independent random variable T.
- Given observations (T_i, Y_i), find regression function: $Y = f(T)$.
- Often phrased as conditional expectation $E[Y|T = t] = f(t)$.
- Parametric (e.g., linear) or nonparametric (e.g., kernel).

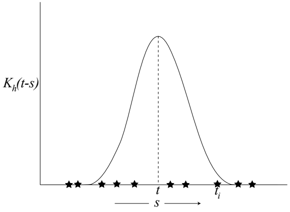
Kernel Regression (Nadaraya-Watson)

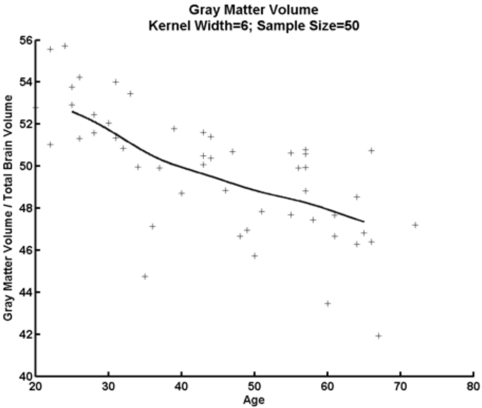
Define regression function through weighted averaging:

$$f(t) = \sum_{i=1}^{N} w_i(t) Y_i$$

$$w_i(t) = \frac{K_h(t - T_i)}{\sum_{i=1}^{N} K_h(t - T_i)}$$

Example: Gray Matter Volume

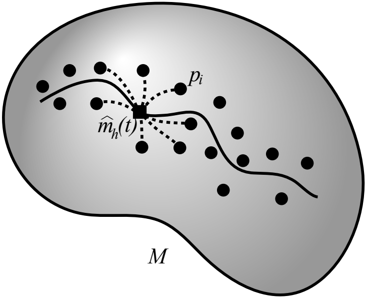




$$w_i(t) = \frac{K_h(t - T_i)}{\sum_{i=1}^{N} K_h(t - T_i)}$$

$$f(t) = \sum_{i=1}^{N} w_i(t) Y_i$$

Manifold Kernel Regression

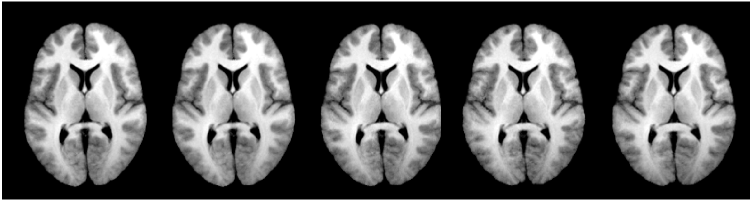


Using Fréchet weighted average:

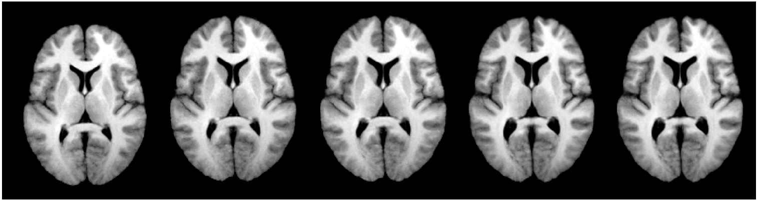
$$\hat{m}_h(t) = \arg\min_y \sum_{i=1}^{N} w_i(t) d(y, Y_i)^2$$

Davis, et al. ICCV 2007

Brain Shape Regression



35 37 39 41 43



45 47 49 51 53

Acknowledgements

Collaborators:

University of Utah

- Sarang Joshi
- Ross Whitaker
- Josh Cates
- Suresh Venkatasubramanian

- Steve Pizer (UNC)
- Brad Davis (Kitware)

Funding:

- NA-MIC, NIH U54 EB005149
- NIH R01 EB007688-01A1

Books

Dryden and Mardia, *Statistical Shape Analysis*, Wiley, 1998.

Small, *The Statistical Theory of Shape*, Springer-Verlag, 1996.

Kendall, Barden and Carne, *Shape and Shape Theory*, Wiley, 1999.

Krim and Yezzi, *Statistics and Analysis of Shapes*, Birkhauser, 2006.

Papers

Kendall, Shape manifolds, Procrustean metrics, and complex projective spaces. *Bull. London Math. Soc.*, 16:18–121, 1984.

Fletcher, Joshi, Lu, Pizer, Principal geodesic analysis for the study of nonlinear statistics of shape, *IEEE TMI*, 23(8):995–1005, 2004.

Pennec, Intrinsic statistics on Riemannian manifolds: Basic Tools for Geometric Measurements. *JMIV*, 25(1):127-154, 2006.

Davis, Fletcher, Bullitt, Joshi. Population shape regression from random design data, ICCV 2007.

Fletcher, Venkatasubramanian, Joshi, The geometric median on Riemannian manifolds with application to robust atlas estimation. *Neuroimage*, 45:S143-52, 2009.