Statistical Methods for Computational Anatomy

Introduction

- Large data sets in which individual samples are represented by shapes (curves and surfaces) are becoming more and more common in medical studies.
- We describe several recent studies, using the tools that were previously described.
- All these methods start with a dataset of *segmented* regions of interest.
- Most address the standard "case vs. controls" statistical problem.

Compute a shape that is central to the dataset (template)...



... Register each shape to the template



Control Shape Evolution using Velocity Vectors





Optimize over velocity vectors (control)

Laurent Younes

COMPUTING AVERAGES

Karcher Means

• On a Riemannian manifold \mathcal{M} , a Karcher mean of a set $\left\{q^{(k)}\right\}_{k=1}^{n}$ is a minimizer of

$$q \mapsto \sum_{k=1}^{n} d_{\mathcal{M}}(q^{(k)}, q)^2$$

• Not always unique, but the minimized function is convex if all $q^{(k)}$'s are close enough to each other.

Relaxed Version

• Minimize

$$\lambda d_{\mathcal{M}}(q^*,q)^2 + \sum_{k=1}^n d_{\mathcal{M}}(q,\overline{q}^{(k)})^2 + U_k(\overline{q}^{(k)})$$

with respect to q and $\left\{\overline{q}^{(k)}\right\}_{k=1}^{n}$, where U_{k} measures the discrepancy between $\overline{q}^{(k)}$ and $q^{(k)}$ and q^{*} is a hyperparameter.

• Can be formalized as a multi-step optimal control problem.



Application: Cardiac Heart Template

- Dataset: 27 MRI scans of normal hearts.
- Hypertemplate: segmented CT scan.
- Cardiac MRIs have coarse resolution (8mm) along the long axis. They are segmented in a collection of planar curves.
- The hypertemplate is a finely triangulated surface.

End-diastole heart template (I)



Single subject sequence: yellow curves are observed, then registered to a surface

Laurent Younes

End-diastole heart template (II)



Average over 27 subjects

Visualizing Geometric Strain on Template



Phase: 0

Cardiac MRI Displacement Covariance Map (ED)





Posterior view

PCA

Working in the tangent plane: PCA

• Once a template is estimated, exponential coordinates can be computed for all shapes in the training set.



PCA in the tangent space

- Let $\{q^{(1)}, \dots, q^{(n)}\}$ be the dataset.
- Call \overline{q} the average shape (or template).
- Compute minimizing geodesics between \overline{q} and each $q^{(k)}$.
- If $d^{(k)}$ is the derivative at time 0 of the geodesic, then

 $q^{(k)} = \exp_{\overline{q}}(d^{(k)})$

• Use $d^{(1)}, \dots, d^{(n)}$ as a new dataset.

PCA in the co-tangent space

- For objects represented as point sets, using the optimal control approach, each $d^{(j)}$ can be replaced by a co-state $p^{(j)}$, which is more convenient.
- The metric used for PCA should be inherited by the Riemannian structure, namely

$$\left\langle p^{(i)}, p^{(j)} \right\rangle_{\overline{q}} = \sum_{k,l=1}^{N} (p^{(i)}_k)^T K_V(\overline{q}_k, \overline{q}_l) p^{(j)}_l$$

(rather than $\sum_{k=1}^{N} (p^{(i)}_k)^T p^{(j)}_k$). Here, $(\overline{q}_1, \dots, \overline{q}_N)$ are 3D

points that form the template.

,

PCA in the co-tangent space

• PCA now finds a basis (e^1, \dots, e^m) , with $e^j = (e_1^j, \dots, e_N^j)$ that is orthonormal for $\langle \cdot, \cdot \rangle_{\overline{a}}$ and minimizes

$$\sum_{j=1}^{n} \left\| p^{(j)} - \pi_m(p^{(j)}) \right\|_{\bar{q}}^2$$

where π_m is the orthonormal projection on span (e^1, \ldots, e^m)

• Each decomposition $p = \sum_{s=1}^{\infty} \alpha_s e^s$ represents a new shape via the exponential map. s=1

Example: BIOCARD dataset projected on first two principal components



Application: Active Shape Models

- PCA provides a prior distribution on shapes around a given template, that can be used to drive or regularize segmentation algorithms.
- This leads to what we call GCDAS (for geodesicallycontrolled diffeomorphic active shapes).
- They solve, for some function U:

$$\frac{1}{2}\sum_{s=1}^{m}\frac{\alpha_{s}^{2}}{\lambda_{s}^{2}}+U(\exp_{\overline{q}}(p))\rightarrow\min$$

with $p = \sum_{s=1}^{m} \alpha_s e^s$ and $(\lambda_1^2, \dots, \lambda_m^2)$ are the eigenvalues of the PCA decomposition.

Example: Landmark extrapolation

- Assume that landmarks have been placed on a target shape.
- The goal is to find a deformation that moves homologous landmarks on the template towards their targets.
- Using GC-DAS, the deformation is also constrained by the prior.

Landmark Matching



Without GC-DAS



With GC-DAS

DIFFEOMORPHOMETRY

Shape Markers

- In addition to the exponential chart coordinates (*p*), other shape descriptors can be used in statistical studies.
- They can be derived from the fact that computing geodesics also provides a diffeomorphic registration of the target shape to the template.
- So, from $(q^{(1)}, \dots, q^{(n)})$, one computes mappings $(\varphi^{(1)}, \dots, \varphi^{(n)})$ which are such that $q^{(k)} \simeq \varphi^{(k)}(\overline{q})$.
- (Diffeo)morphometric markers are based on these diffeomorphisms.

Examples of Shape Markers

- The Jacobian determinant is $J^{\varphi} = \det(D\varphi)$ where $D\varphi$ is the matrix of partial derivatives of φ .
- It measures infinitesimal changes in volumes $\operatorname{vol}(\varphi(\delta v)) \simeq J^{\varphi}(x)\operatorname{vol}(\delta v)$ where δv is a small neighborhood of x.
- Directional changes can also be used. If |u|=1, $|D\varphi(x)u-u|^2$ measures the strain at x in the direction u.
- Eigenvector and eigenvalues of the strain tensor $(D\varphi(x)-I)^T(D\varphi(x)-I)$ can also be used

Case of Surfaces

- When comparing surfaces, the transformation of areas rather than volumes can more relevant.
- If *S* is a surface and $x \in S$, define $J_S^{\varphi}(x)$ so that $\operatorname{area}(\varphi(\delta s)) \simeq J_S^{\varphi}(x)\operatorname{area}(\delta s)$ where δs is a small surface element around *x*.
- It is given by $J_S^{\varphi} = \left\| D \varphi^{-T} N \right\|$ where *N* is the unit normal to *S* at *x*.
- It is easily computable from triangulated surfaces.

Dimension Reduction and Aggregation

- Volumes can have millions of voxels, and surfaces thousands of vertices.
- Although data can be handled at this level, it is sometimes useful to reduce their dimension.
- PCA, applied to the shape markers, is always an option.
- On surfaces, one can also use geometrically induced orthonormal families.

Laplace-Beltrami Eigenvectors

• If *S* is a surface, the Laplace-Beltrami (LB) operator Δ is defined by the identity

$$\int_{S} \nabla f \cdot \nabla g \, ds = -\int_{S} f \, \Delta g \, ds$$

where ∇ is the gradient operator on *S*, i.e., the usual gradient projected on the tangent plane.

• It is a non-positive symmetric operator. Its eigenvectors form an orthonormal basis of $L^2(S)$, called surface harmonics.

Examples



Spectral Segmentation

- One can also reduce dimension by averaging over sub-regions.
- These regions can sometimes be given a priori, as a segmented atlas.
- One can also use geometrically-induced regions, such as those obtained via spectral segmentation using LB eigenvectors.
- These methods attach to each point in the surface the value of the first *m* eigenvectors of the LB operator at this point.
- This is followed by standard clustering (e.g., *K*-means), based on this *m*-dimensional feature.

Illustration







APPLICATION: PREDICT-HD

Geilo, January 2014

Laurent Younes

Predict HD

- Huntington Disease (HD) is caused by a CAG repeat expansion in the huntingtin gene. Individuals with longer repeats develop HD with earlier ages of onset. Predictors highly correlated to the time to onset can be developed (combining age and CAG repeat length).
- HD involves preferential atrophy of the striatal complex (caudate, putamen, nucleus accumbens) and related subcortical nuclei.
- Dataset with 80 subjects belonging to four groups: *controls*, *low*, *mid* and *high* (all prodromal subjects) with labels based on CAG repeat and age (which are combined in a *CAP score*).
- For each subject, segmented substructures were computed (in the form of binary volumes), for left and right accumbens, caudate, globus pallidus, hippocampus, putamen and thalamus.
- Goal: relate shape of structures to groups.

Population Statistics

- 80 subjects in 4 groups
- MRI scan for each subject, with segmented accumbens, caudate, globus pallidus, hippocampus, putamen and thalamus.

	Controls	Low	Mid	High	Total
Female/Male	10/10	10/6	11/7	17/9	48/32
Average age (standard deviation)	42.8 (10.6)	36.7 (8.9)	41.7 (9.1)	43.3 (8.7)	41.5 (9.5)
Population Templates



Shape Markers: surface Jacobian determinants (examples)



Statistical Model

- Let J_k^j denote the *k*th coordinate of the deformation marker for subject *j*.
- Use covariates for gender, age and intracranial volume.
- Let y_k be the residual of the linear regression of J_k by the covariates.
- Let g_{low} , g_{mid} and g_{high} denote binary variables for the low, mid and high CAP score groups (all null for controls)

Three linear models models with null

• Low only:

$$y_{k}^{j} = b_{k,0} + b_{k,\text{low}} g_{\text{low}}^{j} + n_{k}^{j}$$
$$H_{0}(k): b_{k,\text{low}} = 0$$

• Mid and Low:

$$y_{k}^{j} = b_{k,0} + b_{k,\text{low}} g_{\text{low}}^{j} + b_{k,\text{mid}} g_{\text{mid}}^{j} + n_{k}^{j}$$
$$H_{0}(k): b_{k,\text{low}} = b_{k,\text{mid}} = 0$$

• All groups

$$y_{k}^{j} = b_{k,0} + b_{k,\text{low}}g_{\text{low}}^{j} + b_{k,\text{mid}}g_{\text{mid}}^{j} + b_{k,\text{high}}g_{\text{high}}^{j} + n_{k}^{j}$$
$$H_{0}(k): b_{k,\text{low}} = b_{k,\text{mid}} = b_{k,\text{high}} = 0$$

P values (group comparisons, volumes)

Substructure	Low	Mid	High
Accumbens (Left)	0.01	0.09	0.025
Accumbens (Right)	0.11	0.40	0.18
Caudate (Left)	0.62	0.039	< 0.0001
Caudate (Right)	0.55	0.10	< 0.0001
Globus Pallidus (Left)	0.05	0.0003	< 0.0001
Globus Pallidus (Right)	0.002	0.0003	< 0.0001
Hippocampus (Left)	0.77	0.82	0.58
Hippocampus (Right)	0.32	0.62	0.30
Putamen (Left)	0.57	0.0007	< 0.0001
Putamen (Right)	0.41	0.037	< 0.0001
Thalamus (Left)	0.60	0.51	0.23
Thalamus (Right)	0.79	0.90	0.67

P values (group comparisons, shape)

Substructure	Low	Mid	High
Accumbens (Left)	0.44	0.25	0.064
Accumbens (Right)	0.37	0.30	0.035
Caudate (Left)	0.91	0.025	< 0.0001
Caudate (Right)	0.08	0.14	< 0.0001
Globus Pallidus (Left)	0.02	0.001	< 0.0001
Globus Pallidus (Right)	0.0042	0.014	< 0.0001
Hippocampus (Left)	0.18	0.23	0.18
Hippocampus (Right)	0.64	0.74	0.23
Putamen (Left)	0.33	<0.0001	< 0.0001
Putamen (Right)	0.011	0.018	< 0.0001
Thalamus (Left)	0.13	0.59	0.17
Thalamus (Right)	0.50	0.66	0.04

Atrophy maps: Caudate



Left/Right caudate, high CAP group

Geilo, January 2014

Laurent Younes

Putamen (left/right, high CAP)



Left Globus



Right globus







Caudate (rank based)



Putamen (Rank based)



Globus (rank-based)





BIOCARD

Longitudinal Example: BIOCARD (NIH, JHU)

- 1995-2005: Alzheimer's disease longitudinal study at NIH
- Subjects: around 350 healthy subjects, with a large proportion at risk of dementia.
- Study was extended in 2008 (M. Albert) with updated diagnoses (2010) with more subjects converting to the disease.
- MRI data was acquired multiple times for each subject during the first study (1 to six scans per subject).
- Goal: identify shape structures that are primarily affected.

BIOCARD (continued)

- All subjects were "normal" at the beginning of the study (1995).
- At the end of the first study, a small number (15) were diagnosed with mild cognitive impairment (MCI) or dementia (DAT).
- Recent diagnoses in the extended study (2010) reveled 51 patients that converted from controls to cognitively impaired.

General model

- Variables:
 - $\mathcal{Y}_{k}^{j,q}$: *k*th shape marker for subject *j* from scan *q* (corrected for gender and intracranial volume).
 - $t^{j,q}$: age of subject *j* at scan *q*.
 - g^{j} : group of subject *j*.
- Model:

$$y_{k}^{j,q} = a_{k} + a_{k}'t^{j,q} + (b_{k} + \beta b_{k}'t^{j,q})g^{j} + n_{k}^{j,q}$$

• Null hypothesis:

$$b_k = b'_k = 0$$

Noise model (random effects)

• The noise is modeled as

$$n_k^{j,q} = \varepsilon_k^j + \eta_k^{j,q}$$

with $\mathcal{E}_k^j \sim \mathcal{N}(0, rs_k^2)$ and $\eta_k^{j,q} \sim \mathcal{N}(0, s_k^2)$.

- Parameters $(a_k, a'_k, b_k, b'_k, s^2_k, k = 1, ..., d)$ and *r* are estimated by maximum likelihood.
- (Note: *r* is chosen independent from *k* for simplicity and computational efficiency).

• The test statistic is the log-likelihood difference between the null hypothesis $H_0^k : b_k = b'_k = 0$ and the general hypothesis H_1^k :

$$S_{k} = L_{H_{1}}^{k} - L_{H_{0}}^{k}$$

• The log-likelihood in each case is given by

$$-2L^{k} = \operatorname{cst} + N_{\text{subj}} \log \hat{s}_{k}^{2} + \sum_{j} \log(\hat{r}N^{j} + 1)$$

where N_{subj} is the total number of subjects and N^{j} is the number of observations (scans) for subject *j*.

• A global statistic can then be defined by

$$S^* = \max_k S^k$$

- P-values are computed using permutation sampling (scramble groups...) run until a 10% accuracy is reached with high probability.
- Variables y_k for which S^k is larger that the 95 percentile of the values of S^* observed via permutations are considered as significant *at 5% family-wise error rate*.

Significant P-values controls vs. preclinicals

• Strong significance for ERC

Structures	Vertex Laplace		Volume	
Examined	Controls vs.	Controls vs.	Controls vs.	
	Preclinical Preclinical		Preclinical	
	AD AD		AD	
Amygdala (L)	0.17	0.13	0.0086	
Hippocampus (L)	0.022	0.33	0.073	
ERC (L)	<0.0001	0.0001	0.51	
Amygdala (R)	0.031	0.029	0.0043	
Hippocampus (R)	0.025	0.08	0.79	
ERC (R)	0.0067	0.0003	0.17	

Detected Regions





ADNI (CASE-CONTROL)

Laurent Younes

Description of the Study

- The ADNI study was launched in 2003 by a conglomerate of federal agencies, private pharmaceutical companies and non-profit organizations.
- The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD).
- Included subjects: 210 HC, 369 with MCI, and 175 with AD.
- MCI subgroups: MCI-stable reverted to normal cognitive status or remained stable (205); and MCI-AD converted to AD (151 subjects), after one year. (13 reverted to control and were excluded).



• 7 subcortical structures were considered: amygdala, hippocampus, caudate, putamen, thalamus, globus pallidus and ventricle (all left and right).



P value tables

	HC VS. AD		HC VS. MCI		MCI VS. AD	
	volume	shape	volume	shape	volume	shape
	analysis	analysis	analysis	analysis	analysis	analysis
Left-amyg	p < 1E-4	p < 1E-4	p < 1E-4	p < 1.4E-4	p < 1E-4	p < 1E-4
Right-amyg	p < 1E-4	p < 1E-4	p < 1E-4	p < 1E-4	p < 1E-4	p < 1E-4
Left-hipp	p < 1E-4	p < 1E-4	p < 1E-4	p < 1E-4	p < 1E-4	p < 1E-4
Right-hipp	p < 1E-4	p < 1E-4	p < 1E-4	p < 2.6E-4	p < 1E-4	p < 1E-4
Left-vent	p < 1E-4	p < 1E-4	p < 0.04	p < 1E-4	p < 1E-4	p < 1E-4
Right-vent	p < 1E-4	p < 1E-4	p < 5E-3	p < 3.7E-4	p < 1E-4	p < 1E-4
Left-caud	p < 0.46	p < 1E-4	p < 0.52	p < 0.07	p < 0.15	p < 3.7E-4
Right-caud	p < 0.89	p < 1E-4	p < 0.19	p < 0.01	p < 0.17	p < 0.12
Left-puta	p < 0.02	p < 1E-4	p < 0.47	p < 1E-4	p < 0.05	p < 5.9E-4
Right-puta	p < 0.04	p < 1E-4	p < 0.95	p < 2.6E-4	p < 0.02	p < 5.6E-3
Left-thal	p < 0.07	p < 1E-4	p < 0.96	p < 1.2E-4	p < 0.09	p < 4E-4
Right-thal	p < 0.05	p < 1E-4	p < 0.67	p < 1E-4	p < 0.1	p <1E-3
Left-pall	p < 0.08	p < 1E-4	p < 0.27	p < 0.21	p < 5E-3	p < 1.3E-3
Right-pall	p < 0.45	p < 1E-4	p < 0.12	p < 2E-3	p < 0.02	p < 0.02

Amygdala



Hippocampus



Ventricle



Basal Ganglia



MCI stable vs. MCI HD



MCI-stable vs. MCI-AD



Geilo, January 2014

Laurent Younes

REMARK ON LONGITUDINAL DATA ON MANIFOLDS...



- Assume that subject *j* has n_i observations over time.
- Define the first observation as the baseline.
- Compute a template shape based on baselines.
- Goal: represent all shapes in a template-centered coordinate system

First Method

• Compute (like before) coordinates $y^{(j,q)}, q = 1, ..., n_j$ for subject *j* by computing geodesics between the template and all other shapes

Second Method

- Compute template centered coordinates for baselines and baseline-centered coordinates for follow-ups
- Transport baseline-to-follow-up shape information back to the template



Exponential Charts and Transport

- Let T be the template, B a baseline, F a follow-up
- *B* has exponential coordinates at *T*, *F* has exponential coordinates at *B*, i.e.

 $B = \exp_T(D), F = \exp_B(\Delta)$

- Goal: transport the representation Δ from *B* to *T*, resulting in a new representation Δ' and a new shape $F' = \exp_T(\Delta')$
- Linear analog:

$$B = T + D; F = B + \Delta; \Delta' = \Delta; F' = T + \Delta'$$
Parallel Transport

- In nonlinear spaces, there is, in general, no canonical way for transporting coordinate systems.
- On Riemannian manifolds, this can be done *along a curve* using parallel transport
 - The transformation between coordinates is isometric
 - It generally depends on the chosen curve
 - Its computation requires solving a somewhat complicated dynamical system

Parallel Transport for Point Sets...

$$\sum_{b=1}^{N} K(x_{a}, x_{b})(2\partial_{t}h_{b} + \sum_{c=1}^{N} (r_{b}^{T}h_{c} + h_{b}^{T}r_{c})\nabla_{1}K(x_{b}, x_{c})) = \sum_{b=1}^{n} \nabla_{1}K(x_{a}, x_{b})^{T}v^{a}r_{b} + \sum_{b=1}^{n} \nabla_{1}K(x_{b}, x_{a})^{T}v^{b}r_{b} - \sum_{c=1}^{n} \nabla_{1}K(x_{a}, x_{c})^{T}u^{a}h_{c} - \sum_{c=1}^{n} \nabla_{1}K(x_{c}, x_{a})^{T}u^{c}h_{c}$$

with
$$u^a = \sum_b K(x_a, x_b) r_b$$
 and $v^a = \sum_b K(x_a, x_b) h_b$





HYPERTROPHIC CARDIOMYOPATHY

Surface to curves LDDMM





Population statistics

Population	N (F/M)	LVEF (%)	EDLLV(ml)	ESLLV(ml)	LVM (g)	Mean age
НСМ	9 (5/4)	67.8±7.9	130.0±45.9	42.6±19.8	147.0±19.8	45±12
HHD	11 (2-9)	56.3±7.3	168.2±61.6	76.7±39.4	136.5±39.4	53±10
P-value	N/A	0.008	0.18	0.08	0.94	0.18

End-Systole Shape Analysis (crosssectional)



Radial

FDR

ES-ED transformation (longitudinal)

s I 0.05 0.045 0.04 0.035 0.03 0.025 0.02 0.015 0.01 0.005 ?

Circumferential

FDR

FWER



RISK PREDICTION: LDA ON ADNI

Laurent Younes

Preprocessing

- 7x2 brain structures are analyzed (baselines only).
- For each of them, compute a template, and exponential coordinates at the template.
- Reduce dimension using group-independent PCA.
- Groups: 210 HC, 369 with MCI, and 175 with AD.
- MCI subgroups: 205 MCI-stable, 151 MCI-AD (13 "reverters").

Selecting the optimal LDA classifier procedure

Step1:



Results

- Shape PC information is more discriminating than volume
- Hippocampus is the most discriminant.
- The optimal LDA classifier combines the hippocampus, amygdala and lateral ventricle.
- In the double loop cross-validation, hippocampus was selected 88% of the time, amygdala 83%, ventricle 71%, thalamus 45%, caudate 36%, putamen 37% and pallidum 26%.
- The leave-one-out cross-validation procedure yields correct classification rates: 88% for HC, 86% for AD and 86% for the two groups together.
- Using volumes only: 76% for HC, 75% for AD group, and 75% for the two groups.

Some References (statistical studies)

- Shape abnormalities of subcortical and ventricular structures in mild cognitive impairment and Alzheimer's disease: Detecting, quantifying, and predicting X Tang, D Holland, AM Dale, L Younes, MI Miller Human Brain Mapping 2014
- The diffeomorphometry of temporal lobe structures in preclinical Alzheimer's disease MI Miller, L Younes, JT Ratnanather, T Brown, H Trinh, E Postell, DS Lee, MC ... NeuroImage: Clinical 3, 352-360 2013
- Robust Diffeomorphic Mapping via Geodesically Controlled Active Shapes D Tward, J Ma, M Miller, L Younes, International journal of biomedical imaging 2013
- Principal component based diffeomorphic surface mapping, A Qiu, L Younes, MI Miller, Medical Imaging, IEEE Transactions on 31 (2), 302-311 2012
- Regionally selective atrophy of subcortical structures in prodromal HD as revealed by statistical shape analysis L Younes, JT Ratnanather, T Brown, E Aylward, P Nopoulos, H Johnson, VA ... Human brain mapping 2012
- Matching sparse sets of cardiac image cross-sections using large deformation diffeomorphic metric mapping algorithm: S Ardekani, A Jain, S Jain, TP Abraham, MR Abraham, S Zimmerman, RL Winslow ... Statistical Atlases and Computational Models of the Heart. Imaging and ...

More on Diffeomorphic Shape Analysis

- Shape deformation analysis from the optimal control viewpoint: S Arguillere, E Trélat, A Trouvé, L Younes arXiv preprint1401.0661 2014
- Diffeomorphometry and geodesic positioning systems for human anatomy: MI Miller, L Younes, A Trouvé Technology, 1-9 2013
- Spaces and manifolds of shapes in computer vision: An overview: L Younes, Image and Vision Computing 30 (6), 389-397
- Shape Spaces: A Trouvé, L Younes, Handbook of Mathematical Methods in Imaging, 1309-1362 2011
- Diffeomorphic active contours: F Arrate, JT Ratnanather, L Younes, SIAM journal on imaging sciences 3 (2), 176-198, 2010
- Transport of relational structures in groups of diffeomorphisms: L Younes, A Qiu, RL Winslow, MI Miller, Journal of mathematical imaging and vision 32 (1), 41-56
- Shapes and diffeomorphisms: L Younes, Springer, 2010