ShapeLME: Longitudinal statistical analysis tutorial

Introduction

Deformetrica gave us a way to estimate smooth subject specific shape models. Each subject model lets us re-evaluate shapes at known time-points. Given a longitudinal data set, a more natural question is to understand the behavior of an entire group (i.e., healthy group, groups at risk of HD) and also to have the ability to distinguish between two different groups. The optimization problem for diffeomorphic longitudinal modeling is challenging and existing approaches haven't yet provided a consistent way to compare two different trajectories. But the estimation of diffeomorphic subject specific models remove extraneous and unwanted variability in data.

To do statistics, this package **ShapeLME** computes longitudinal models to summarize the behavior of a specific group. To do longitudinal analysis, we first make sure that the output of Deformetrica processing gives shapes in longitudinal correspondence. With these correspondences, the setup is ripe to estimate linear mixed-effects for shape as summaries [2, 1].

Diffeomorphic trajectories are good at capturing smooth anatomical shape changes, while the particle optimization framework excels at finding compact statistical shape representations with increased statistical power. The idea here is to leverage the strengths of both approaches, to provide an integrated solution, characterized by improved statistical performance in the analysis of both scalar and shape trajectory data derived from noisy segmentations [3].

Installation

- Create and cd into a **build** directory in the **shape-lme** package part of the tutorial.
- Run **ccmake ../**
- Once compiled and built: **cd deformetrica/tutorial/scripts/**

We are now ready to do some statistical analysis.

From shape meshes to point correspondences

Each observation of each subject are estimated and stored in vtk format with information about mesh points, polygonal connectivity surface normals, and every shape observation is evaluated in correspondence with every other shape. We extract just the point correspondence positions alone and store it in another directory in lpts format.

From vtk to lpts

- **data-path**: path to **regressed observed shape meshes** - our naming: reg_at_obs_time_pts
- **output-data-path**: path to output point correspondence files - our naming: shape_lpts_for_stats

Run the script:

- **python write_vtk_to_lpts_script.py**

Estimating longitudinal trajectories

The linear mixed-effects model, originally developed by Laird and Ware [2] has the ability to incorporate multiple covariates such as risk of onset group membership (CTRL, LOW, MED, HIGH), sex, among other information and estimate an associated model. This model was extended to study longitudinal shape by Datar, et al. [1].

For the tutorial, we control for sex by considering only female subjects and estimate models for CTRL and HIGH risk groups only. Once the script is run, output directories for each structure (left/right caudate/putamen) is created and fixed and random effects are respectively estimated and stored as a text file.

Note: Deformetrica estimates diffeomorphic models for entire shape complexes (such as the strital complex of left/right caudate/putamen), whereas here we separate out each structure, center them, align each structure to a template structure, and estimate longitudinal models per structure. This doesn't have to be the case. Here too, one could consider the whole shape complex, with an appropriate alignment, and estimate models that way as long as there is longitudinal correspondence across subjects and time-points. This will be incorporated in future iterations of this software, but for now, we estimate models for each structure separately.

Estimating mixed-effects

- **input-data-path**: the lpts directory path obtained as output above - shape_lpts_for_stats
- **output-data-path**: path to fixed and random effects files for each structure - mixed_effects_results

Run the script:

- **python estimate_mixed_effects_script.py**

Visualizing shapes and trajectories

The baseline shapes along with the respective group evolutions of each category can be visualized in Paraview. To do this, we convert the respective point and velocity files back to vtk format keeping the original neighborhood and connectivity information of the input mesh data. Once thats done, we then load the respective vtk files into Paraview and compare the visualizations there.

- **input-path**: the mixed effects results directory
- **output-data-path**: output vtk shape sequences path for each structure - our naming: output_shape_seq_vtk

Run the script:

- **python create_shape_sequences_script.py**

Hypothesis testing

We finally test the hypothesis for significant differences between CTRL and HIGH risk groups by a non-parametric permutation test using the Hotelling t^2 statistic. The input to this are the point correspondence files, and the output is a p-value.

- **input-data-path**: the lpts directory path obtained as output above - shape_lpts_for_stats
- **output-data-path**: text file listing hypothesis results - we store it in mixed_effects_results

Run the script:

- **python group_hypothesis_testing_script.py**

References

[1] M. Datar, P. Muralidharan, A. Kumar, S. Gouttard, J. Piven, G. Gerig, R.T. Whitaker, and P.T. Fletcher. Mixed-effects shape models for estimating longitudinal changes in anatomy. In *MICCAI STIA*, volume 7570, pages 76–87, 2012.

[2] Nan Laird and James H. Ware. Random-effects models for longitudinal data. *Biometrics*, 38(4):963–974, 1982.

[3] P. Muralidharan, J. Fishbaugh, H.J. Johnson, S. Durrleman, J.S. Paulsen, G. Gerig, and P.T. Fletcher. Diffeomorphic shape trajectories for improved longitudinal segmentation and statistics. In *Proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 2014.