BAYESIAN COVARIATE SELECTION IN MIXED-EFFECTS MODELS FOR LONGITUDINAL SHAPE ANALYSIS

Prasanna Muralidharan[1] *James Fishbaugh*[2] *Eun Young Kim*[3]
Hans J. Johnson[3] *Jane S. Paulsen*[3] *Guido Gerig*[2] *P. Thomas Fletcher*[1]

[1] School of Computing & SCI Institute, University of Utah, [2] CSE, New York University,
[3] Department of Psychiatry, Carver College of Medicine, University of Iowa

ABSTRACT

The goal of longitudinal shape analysis is to understand how anatomical shape changes over time, in response to biological processes, including growth, aging, or disease. In many imaging studies, it is also critical to understand how these shape changes are affected by other factors, such as sex, disease diagnosis, IQ, etc. Current approaches to longitudinal shape analysis have focused on modeling age-related shape changes, but have not included the ability to handle covariates. In this paper, we present a novel Bayesian mixed-effects shape model that incorporates simultaneous relationships between longitudinal shape data and multiple predictors or covariates to the model. Moreover, we place an Automatic Relevance Determination (ARD) prior on the parameters, that lets us automatically select which covariates are most relevant to the model based on observed data. We evaluate our proposed model and inference procedure on a longitudinal study of Huntington's disease from PREDICT-HD. We first show the utility of the ARD prior for model selection in a univariate modeling of striatal volume, and next we apply the full high-dimensional longitudinal shape model to putamen shapes.

Index Terms— Longitudinal shape analysis, Bayesian analysis, model selection, Huntington's disease

1. INTRODUCTION

Longitudinal imaging studies involve tracking subjects by repeated image acquisition over time. A primary goal of longitudinal neuroimaging studies is to find sensitive biomarkers that correlate with disease outcomes. In neurodegenerative diseases, such as Alzheimer's and Huntington's disease, the shape of the brain is affected. Longitudinal statistical shape analysis involves understanding and quantifying anatomical shape variability within and across subjects, its correlation to predictors such as age, clinical scores related to disease, and also to distinguish between normal and disease populations.

Previous work on longitudinal shape modeling includes the use of diffeomorphic mappings by Qiu et al. [1] to track changes in a subject and map the individual trends to a population atlas via parallel transport. Durrleman et al. [2] construct spatiotemporal image atlases from longitudinal data. Barry et al. [3] build mixed-effects models on a small number of manually selected landmarks to model the development of facial shape. Datar et al. [4] build linear mixed-effects models treating shape as a collection of point distribution models in correspondence across subjects. Muralidharan and Fletcher [5] develop a manifold version of a mixed-effects model to analyze longitudinal data taking values on a finite-dimensional Riemannian manifold. Singh et al. [6] develop hierarchical geodesic models in the infinite-dimensional space of diffeomorphisms to study longitudinal imaging data.

All of the above approaches are limited to modeling longitudinal shape as a function of time as the only predictor. However, we usually have a lot more information about subjects in a longitudinal imaging study, such as sex, IQ, diagnosis groups, clinical scores associated to disease, etc. Developing statistical shape models that can handle such covariates is critical for two reasons. First, statistical analysis can often be improved by controlling for nuisance variables, i.e., variables that are not of primary interest but have a significant effect on the model. Second, including categorical variables, such as sex or diagnosis, can help explain how longitudinal shape trends are different in different populations. However, there are also dangers to including covariates in a statistical model. One such danger is that different combinations of covariates can lead to drastic changes in the statistical significance of the variables of interest. This opens the risk of "p-value fishing", where several covariate combinations are attempted in search of the desired result. Another danger of including covariates is that each new covariate adds a number of parameters proportional to the dimension of the response variable (which is

This research was supported by NIH grants U01 NS082086, NS40068, NS050568 (PREDICTHD), U54 EB005149 (NA-MIC), S10 RR023392 (NCCR Shared Instrumentation Grant), NSF CAREER grant 1054057, NIH (NINDS; 5RO1NS040068, 5RO1NS054893) and the CHDI Foundation to Jane S Paulsen. We thank the PREDICT-HD sites, the study participants the National Research Roster for HD patients and Families, the Huntington Disease Society of America and the Huntington Study Group.

very large in the case of shape responses).

To this end, we present a novel Bayesian mixed-effects model for longitudinal shape data that incorporates relationships between shape change and multiple predictors simultaneously. Our first principle is that the model should automatically choose the appropriate covariates to include in a data-driven fashion, avoiding the need for ad hoc choices from the user. The second principle is Occam's razor, that the model should be no more complex than is needed to explain the data. To this end, our model uses an automatic relevance determination (ARD) prior on all fixed-effects covariates, which drives the irrelevant coefficients to zero that do not have a significant contribution to the model given the data.

2. THE MODEL

2.1. Background on Mixed-Effects Models

In a linear mixed-effects (LME) model [7], the response or observed variable y_i is assumed to have a set of p parameters α, fixed across n subjects, representative of population parameters, called "fixed-effects". In addition, the i^{th} subject is assigned a vector of q subject-specific parameters, β_i, called random-effects that model the deviation of the subject from the population. For $i \in \{1, 2, \ldots, n\}$, the LME model is

$$y_i = X_i\alpha + Z_i\beta_i + \epsilon_i, \tag{1}$$

where for the i^{th} individual, X_i and Z_i are known design matrices with covariate information that influence y_i through fixed and random effects respectively, $\beta_i \sim N(0, D)$, with arbitrary covariance matrix D, and $\epsilon_i \sim N(0, \tau^{-1}I_i)$, for some precision τ and I_i identity matrix.

2.2. Univariate LME with automatic covariate selection

For a univariate longitudinal response variable y with fixed-effects α and random-effects β, Armagan et al. [8] propose the following Bayesian model:

Likelihood:

$$p(y|\alpha, \beta; \tau) = \left(\frac{\tau}{2\pi}\right)^{\frac{N}{2}} \exp\left(-\frac{\tau}{2}\sum_{i=1}^{n}\|y_i - X_i\alpha - Z_i\beta_i\|^2\right), \tag{2}$$

where N denotes the total number of observations (all time-points of all subjects).

Prior on α: To select covariates that are most relevant to the model, we place an automatic relevance determination (ARD) prior on α [9]: $\alpha \sim N(0, \Omega^{-1})$, where $\Omega = \text{diag}(\omega_k), k = 1, \ldots, p$ is a diagonal matrix of Gaussian precision parameters, i.e.,

$$p(\alpha|\Omega) = \frac{(\omega_1 \ldots \omega_p)^{1/2}}{2\pi^{p/2}} \exp\left(-\frac{1}{2}\alpha'\Omega\alpha\right). \tag{3}$$

Posterior for α:

$$p(\alpha|y, \beta, \tau, \Omega) \propto p(y|\alpha, \beta, \tau)p(\alpha|\Omega) \sim N(\hat{\alpha}, \hat{A}), \tag{4}$$

where

$$\hat{A} = (\Omega + \tau\sum_{i=1}^{n} X_i'X_i)^{-1}, \quad \hat{\alpha} = \tau\hat{A}\sum_{i=1}^{n} X_i'(y_i - Z_i\beta_i).$$

Inference: In the general case, τ and D are unknown parameters, and [7] develop an expectation maximization (EM) algorithm to compute restricted maximum likelihood (REML) estimates of all model parameters including α and $\beta_i, i = 1, \ldots, n$ in an iterative manner. In contrast with [7], α here is a random variable with the chosen ARD prior with precision parameters $\omega_k, k = 1, \ldots, p$. From (4), see that the posterior distribution of α is Gaussian. For Bayesian inference, we take a similar EM approach as in [7], but with the added estimation of ω's and computing the MAP estimate of α in every EM iteration instead of the REML estimate. On convergence of the EM algorithm, a high ω estimate implies that the posterior distribution of the associated covariate will peak about 0, and hence be deemed irrelevant to the model, whereas a low ω value keeps the associated covariate in the model. See [8] for a treatment of univariate Bayesian mixed-effects models that also includes prior choices for random effects.

2.3. Shape LME with automatic covariate selection

We now propose a novel Bayesian mixed-effects model to study longitudinal shape evolution that automatically selects relevant covariates associated to shape change. As in [4], shapes are represented as point distributions in correspondence across subjects and time-points. The joint shape likelihood can be written as

$$p(y|\alpha, \beta; \tau) = \frac{\prod_{l=1}^{m}\tau_l^{\frac{N}{2}}}{(2\pi)^{\frac{mN}{2}}} \exp\left(-\frac{1}{2}\sum_{l=1}^{m}\tau_l\sum_{i=1}^{n}\|y_i^l - X_i\alpha^l - Z_i\beta_i^l\|^2\right), \tag{5}$$

where y is a shape response variable with m independent co-ordinates, and α, β are now fixed and random effects "shape" parameters. In the particle representation of shapes, note that each co-ordinate of each particle being univariate longitudinal, follow their own independent 1D LME model with unknown parameters τ^l, D^l. We place an ARD prior on fixed effects α, given as

$$p(\alpha|\Omega) = \prod_{l=1}^{m} p(\alpha^l|\Omega) = \frac{(\omega_1 \ldots \omega_p)^{m/2}}{2\pi^{mp/2}} \exp\left(-\frac{1}{2}\sum_{l=1}^{m}(\alpha^l)'\Omega\alpha^l\right), \tag{6}$$

where Ω is a diagonal covariance matrix. Note that we choose Ω to be shared across all coordinates of all particles representing the shape. We could instead model separate covariances for each coordinate. However, a shared Ω allows us to select relevant covariates which influence the shape as a whole. For example, we can ask if sex is a relevant predictor for global shape.

Inference: Unlike the univariate case, the posterior distribution of α isn't Gaussian but rather an m-product of Gaussians. For inference, we follow a similar EM procedure to that in Section 2.2, with the difference being that Ω estimation depends on the entire shape corpus, i.e., the current MAP estimate of fixed-effects α and estimates of all $\tau^l, l = 1, \ldots, m$. The estimate for $\omega_k, k = 1, \ldots, p$, denoted $\hat{\omega_k}$ maximizes $p(y|\Omega, \beta, \tau)$, and we derive this as closed form solution.

$$\begin{aligned}\hat{\omega_k} &= \arg_\omega \max p(y|\Omega, \beta, \tau) \\ &= \arg_\omega \max \prod_{l=1}^{m} \int_{\alpha^l} p(y^l|\alpha^l, \beta^l, \tau^l) p(\alpha^l|\Omega) d\alpha^l \\ &= \frac{\sum_{l=1}^{m} \mathrm{tr}(\hat{A}^l E_{kk} \sum_{i=1}^{n} X_i' X_i)}{\sum_{l=1}^{m} \hat{\alpha^l}(k)^2}, \end{aligned} \tag{7}$$

where $\hat{A}^l = (\tau^l \Omega + \sum_{l=1}^{n} X_i' X_i)^{-1}$, E_{kk} is a $p \times p$ that is 1 at the k^{th} diagonal position and zero everywhere else.

3. EXPERIMENTS

Data: We study subcortical change associated with Huntingtons disease (HD), leveraging the longitudinal study PREDICT-HD. The longitudinal database consists of 209 female subjects (66 CTRL, 36 LOW, 41 MED, 66 HIGH) and 112 male subjects (42 CTRL, 13 LOW, 14 MED, 43 HIGH). The LOW / MED / HIGH categories represent probability of onset of manifesting signs of HD. All subjects have had at least 2 MR images acquired approximately one year apart, with many subjects undergoing multiple scans per visit. Six subcortical pairs (caudate, putamen, hippocampus, thalamus, acumben, and pallidus) were segmented from each MR image. For our experiments in this paper, we restrict our attention to the striatal complex (left/right caudate and putamen).

Preprocessing: The neurodegeneration process associated with HD has been observed as a temporally smooth process [10]. An emerging model of smooth anatomical change is to consider continuous transformations of the ambient space by differentiable and invertible deformations. To remove extraneous variability from raw imaging data, we estimate continuous and temporally consistent sequence of shapes as prescribed by the first stage of consistent longitudinal segmentation in [11]. We then extract longitudinal shape correspondences from these smooth meshes computed at observed time-points. These shape correspondences feed into our Bayesian covariate mixed-effects shape model. We also compute structural volumes of these temporally consistent shapes as a derived measure and use this as our 1D longitudinal data for covariate statistical analysis. Note that the statistical shape model we propose is independent of the way we obtain shape correspondences. The model is applicable to a different set of valid and consistent shape correspondences.

Model: The *full model* with all covariates of interest for longitudinal shape and volume is

$$y = \text{intercept} + \text{age} + \text{sex} + \text{group} + \text{age} \times \text{sex} + \text{age} \times \text{group}. \tag{8}$$

The model accounts for 10 covariates in all since we have 2 sexes (male / female), and 4 groups (CTRL, LOW, MED, HIGH) along with corresponding interaction terms. Our Bayesian inference procedure will select the most "relevant" covariates to the model.

Statistical analysis of Striatal volume: We first evaluate our proposed method to study longitudinal striatal volume change using the model in (8). Our inference procedure selects some of these covariates as most relevant, namely, the intercept, age, sex, age $\times$ LOW, age $\times$ MED and age $\times$ HIGH. Figure 1 shows the estimated covariate model for both males and females. For both sexes, note that the slopes of the risk group lines become more negative as we go from CTRL to LOW to MED to HIGH. Also, see that slopes estimated for males and that for females are the same, since the interaction term of age $\times$ sex was deselected from the model.

Parameter	p-value	Parameter	p-value
age	**0.0088**	age $\times$ LOW	$<$ **0.001**
sex	**0.0223**	age $\times$ MED	$<$ **0.001**
age $\times$ sex	0.9198	age $\times$ HIGH	$<$ **0.001**

Table 1. Fixed-effects covariates significance values estimated from LME analysis of longitudinal striatal volumes

To check if "relevant" covariates we inferred make sense, we also computed Akaike information criterion (AIC) values of all possible 2^{10} models for this data (since there are 10 covariates). Our model was one of 4 models with the lowest AIC. In Table 1, we report significance values for different covariates. See that each of interaction terms age $\times$ group are significant to the model but age $\times$ sex isn't.

Statistical analysis of Longitudinal shape: We next evaluate our method on longitudinal right putamen. The intercept, age, sex, age $\times$ group, and the intercept interaction with MED and HIGH were the selected covariates based on our estimation. The covariates LOW, age $\times$ sex were deselected. The color map (from blue to red) is generated as the dot product of the age parameter with surface normals depicting local volume change. (See Figure 2). The irrelevance of age $\times$ sex is corroborated by a near identical colormap in both sexes. Also, see that there is a gradual increase in shape change as we go from CTRL to LOW to MED to HIGH. We highlight

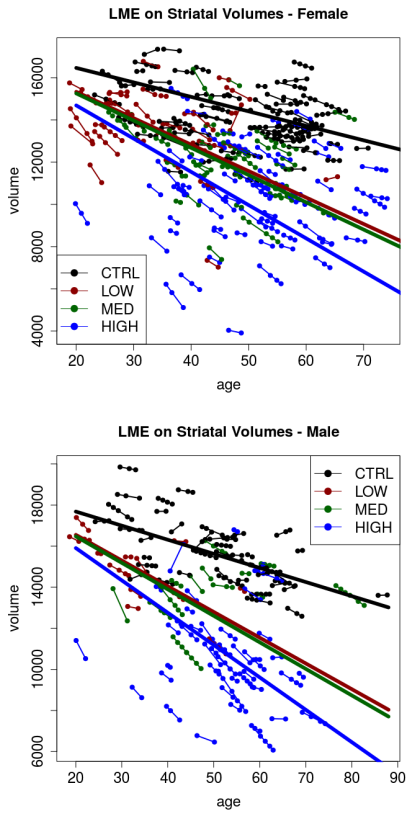

Fig. 1. Bayesian LME covariate analysis on Female (top) and Male (bottom) striatal volumes

that the HIGH category displayed the most amount of twisting and bending when the shape sequence was seen as a 3D evolution. As an additional proof-of-concept experiment, we included a random white noise covariate to the shape model and inferred that this covariate was irrelevant.

4. CONCLUSION

We presented a novel Bayesian method to automatically select relevant covariates that influence global shape change. We evaluated our methods on both longitudinal shape and volume from the PREDICT-HD database. Our shape model assumes independence between random effects of particle coordinates, but can be extended to include spatial correlations. It remains an open question how one can compute Bayesian credible regions and predictive distributions for shape evolutions.

5. REFERENCES

[1] A. Qiu, M. Albert, L. Younes, and M. I. Miller, "Time sequence diffeomorphic metric mapping and parallel transport track time-dependent shape changes," *NeuroImage*, vol. 45, no. 1, 2009.

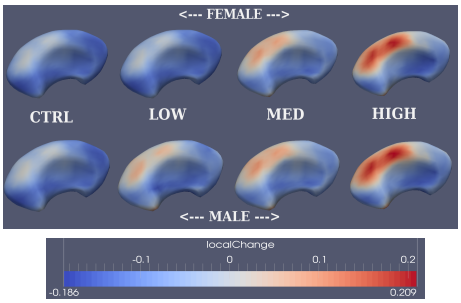

Fig. 2. Longitudinal shape analysis: Right putamen. Colormap: Local contraction (blue) to expansion (red) in mm

[2] S. Durrleman, X. Pennec, A. Trouvé, J. Braga, G. Gerig, and N. Ayache, "Toward a comprehensive framework for the spatiotemporal statistical analysis of longitudinal shape data," *IJCV*, vol. 103, no. 1, 2013.

[3] S. J. E. Barry and A. W. Bowman, "Linear mixed models for longitudinal shape data with applications to facial modeling," *Biostatistics*, vol. 9, no. 3, 2008.

[4] 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*, 2012.

[5] P. Muralidharan and P. T. Fletcher, "Sasaki metrics for the analysis of longitudinal data on manifolds," in *CVPR*, 2012.

[6] N. Singh, J. Hinkle, S. Joshi, and P.T. Fletcher, "Hierarchical geodesic models in diffeomorphisms," *IJCV*, 2015.

[7] N. Laird and J. H. Ware, "Random-effects models for longitudinal data," *Biometrics*, vol. 38, no. 4, 1982.

[8] A. Armagan and D. Dunson, "Sparse variational analysis of linear mixed models for large data sets," *Statistics & probability letters*, vol. 81, no. 8, 2011.

[9] M. E Tipping, "Sparse bayesian learning and the relevance vector machine," *JMLR*, vol. 1, 2001.

[10] E. Aylward, J. Mills, D. Liu, P. Nopoulos, C. A. Ross, R. Pierson, and J. S. Paulsen, "Association between Age and Striatal Volume Stratified by CAG Repeat Length in Prodromal Huntington Disease," *PLoS Curr*, vol. 3, 2011.

[11] 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 *MICCAI*, 2014.