FUNCTIONAL PARTIAL MEMBERSHIP MODELS

Nicholas Marco[1], Damla Şentürk[1], Shafali Jeste [2], Charlotte DiStefano [3], Abigail Dickinson [4], and Donatello Telesca [1]

[1] Department of Biostatistics, University of California, Los Angeles
[2] Division of Neurology and Neurological Institute, Children's Hospital Los Angeles
[3] Division of Psychiatry, Children's Hospital Los Angeles
[4] Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles

Introduction

We extend the class of partial membership models to functional data by defining a partial membership Gaussian process through projections on the linear subspace spanned by a suitable set of basis functions.

Extension of Finite Mixture Models

A typical model-based representation of functional clustering assumes that each underlying process is a latent Gaussian process (GP), $f^{(k)} \sim \mathcal{GP}\left(\mu^{(k)}, C^{(k)}\right)$, with sample paths f_i such that

$$p\left(f_i \mid \rho^{(1:K)}, \mu^{(1:K)}, C^{(1:K)}\right) = \sum_{k=1}^{K} \rho^{(k)}\, \mathcal{GP}\left(f_i \mid \mu^{(k)}, C^{(k)}\right);$$

where $\rho^{(k)}$ is the mixing proportion for component k, s.t. $\sum_{k=1}^K \rho^{(k)} = 1$. Equivalently, by introducing latent variables $\boldsymbol{\pi}_i = [\pi_{i1}\cdots\pi_{iK}]$ ($\pi_{ik} \in \{0,1\}$ and $\sum_{k=1}^K \pi_{ik} = 1$), we have that

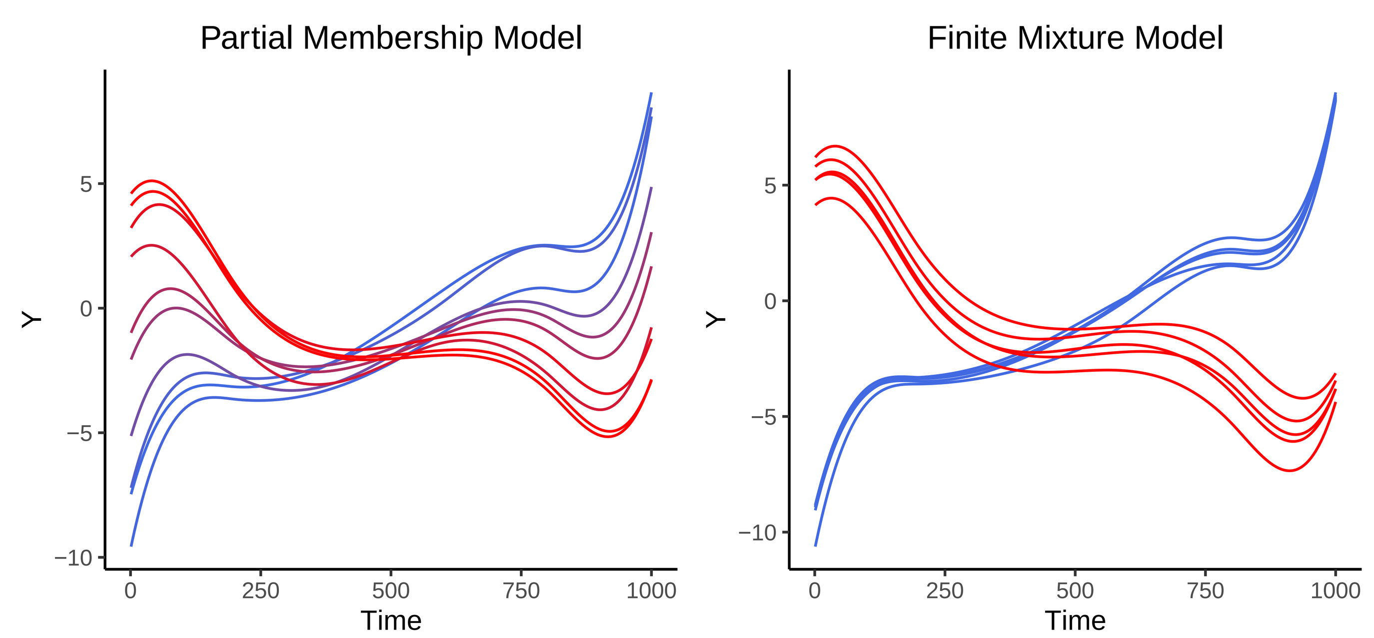
$$p\left(f_i \mid \rho^{(1:K)}, \mu^{(1:K)}, C^{(1:K)}\right) = \sum_{\boldsymbol{\pi}_i} p(\boldsymbol{\pi}_i) \prod_{k=1}^{K} \mathcal{GP}\left(f_i \mid \mu^{(k)}, C^{(k)}\right)^{\pi_{ik}},$$

where $p(\pi_{ik} = 1) = \rho^{(k)}$. By conditioning on the latent π_{ik} variables, we have

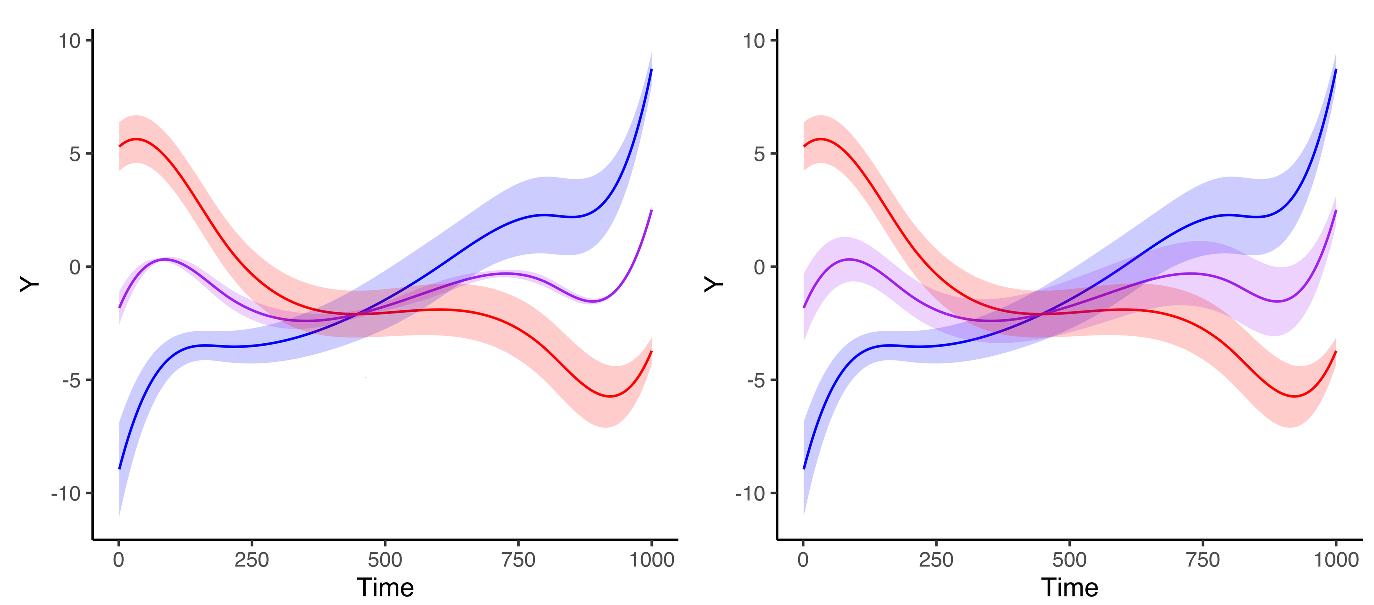
$$f_i \mid \boldsymbol{\pi}_1, \dots, \boldsymbol{\pi}_N = \sum_{k=1}^{K} \pi_{ik} f^{(k)}. \quad (1)$$

By introducing the latent variables $\mathbf{z}_i = [Z_{i1}\cdots Z_{iK}]$, where $Z_{ik} \in (0,1)$ and $\sum_{k=1}^K Z_{ik} = 1$, we arrive at our proposed partial membership model:

$$f_i \mid \mathbf{z}_1, \dots, \mathbf{z}_N = \sum_{k=1}^{K} Z_{ik} f^{(k)}. \quad (2)$$



In equation 2, $f^{(k)}$ can no longer be considered mutually independent. Therefore, we must estimate the cross-covariance functions between the GPs in order to ensure adequately flexible models, as illustrated below.



Functional Partial Membership Models

We will assume $f^{(k)}$ can be represented by a user-defined set of **uniformly continuous** basis function ($\{b_1, \dots, b_P\}$). By assuming that $f^{(k)} \in$ span$(b_1, \dots, b_P)$, we can leverage the Multivariate Karhunen-Loève theorem [2] to get

$$f^{(k)}(t) \approx \boldsymbol{\nu}_k'\mathbf{B}(t) + \sum_{m=1}^{M} \chi_m \boldsymbol{\phi}_{km}'\mathbf{B}(t), \quad (3)$$

with equality when $M = KP$. In equation 3, $\mathbf{B}(t)$ is a vector of the P basis functions evaluated at t, $\boldsymbol{\nu}_k \in \mathbb{R}^P$, $\boldsymbol{\phi}_{km} \in \mathbb{R}^P$, and $\chi_m \sim \mathcal{N}(0,1)$. Using this decomposition, we have that $\mu^{(k)}(t) = \boldsymbol{\nu}_k'\mathbf{B}(t)$ and $C^{(k,j)}(t_k, t_j) \approx \sum_{m=1}^M \boldsymbol{\phi}_{km}'\mathbf{B}(t_k)\boldsymbol{\phi}_{jm}'\mathbf{B}(t_j)$, where $C^{(k,j)}$ is the cross-covariance function between $f^{(k)}$ and $f^{(j)}$. Using this decomposition, we need $\mathcal{O}(KPM)$ parameters to model the covariance structure, while a more naïve method would require $\mathcal{O}(K^2P^2)$ parameters.
Using the decomposition in equation 3, we arrive at our likelihood:

$$y_i(t)|\boldsymbol{\Theta} \sim \mathcal{N}\left(\sum_{k=1}^{K} Z_{ik}\underbrace{\left(\boldsymbol{\nu}_k'\mathbf{B}(t) + \sum_{m=1}^{M}\chi_{im}\boldsymbol{\phi}_{km}'\mathbf{B}(t)\right)}_{f^{(k)}(t)}, \sigma^2\right). \quad (4)$$

Integrating out the χ variables, we arrive at

$$y_i(\mathbf{t}_i)|\boldsymbol{\Theta}_{-\chi} \sim \mathcal{N}\left(\sum_{k=1}^{K} Z_{ik}\underbrace{\mathbf{S}'(\mathbf{t}_i)\boldsymbol{\nu}_k}_{\boldsymbol{\mu}^{(k)}(\mathbf{t}_i)}, \sum_{k=1}^{K}\sum_{j=1}^{K} Z_{ik}Z_{ij}\underbrace{\left(\mathbf{S}'(\mathbf{t}_i)\sum_{m=1}^{M}\left(\boldsymbol{\phi}_{km}\boldsymbol{\phi}_{jm}'\right)\mathbf{S}(\mathbf{t}_i)\right)}_{C^{(k,j)}(\mathbf{t}_i,\mathbf{t}_i)} + \sigma^2\mathbf{I}_{n_i}\right).$$

One of the drawbacks to using the Multivariate Karhunen-Loève theorem is that $\boldsymbol{\phi}_{km}'\mathbf{B}(t)$ construct scaled eigenfunctions of the covariance operator. Therefore, we would normally have to sample from a constrained parameter space to ensure that the eigenfunctions are mutually orthogonal. However, we show that we can still recover all the parameters of interest when relaxing the orthogonality constraint.

Autism Spectrum Disorder (ASD)

Autism spectrum disorder (ASD) is a term used to describe individuals with a collection of social communication deficits and restricted or repetitive sensory-motor behaviors [3]. In the following case studies, we will be analyzing electroencephalogram (EEG) data of 39 typically developing (TD) children and 58 children with ASD between the ages of 2 and 12 years old. Each child was instructed to look at a computer monitor displaying bubbles for two minutes in a dark, sound-attenuated room [1].

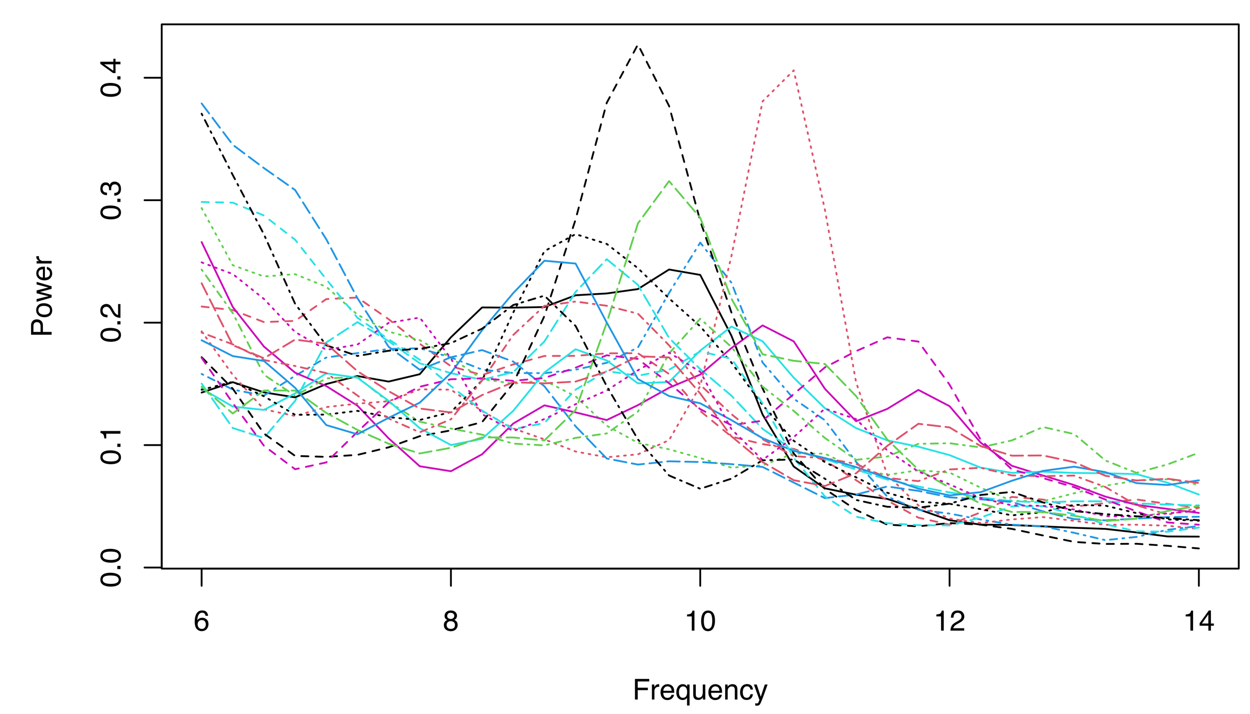
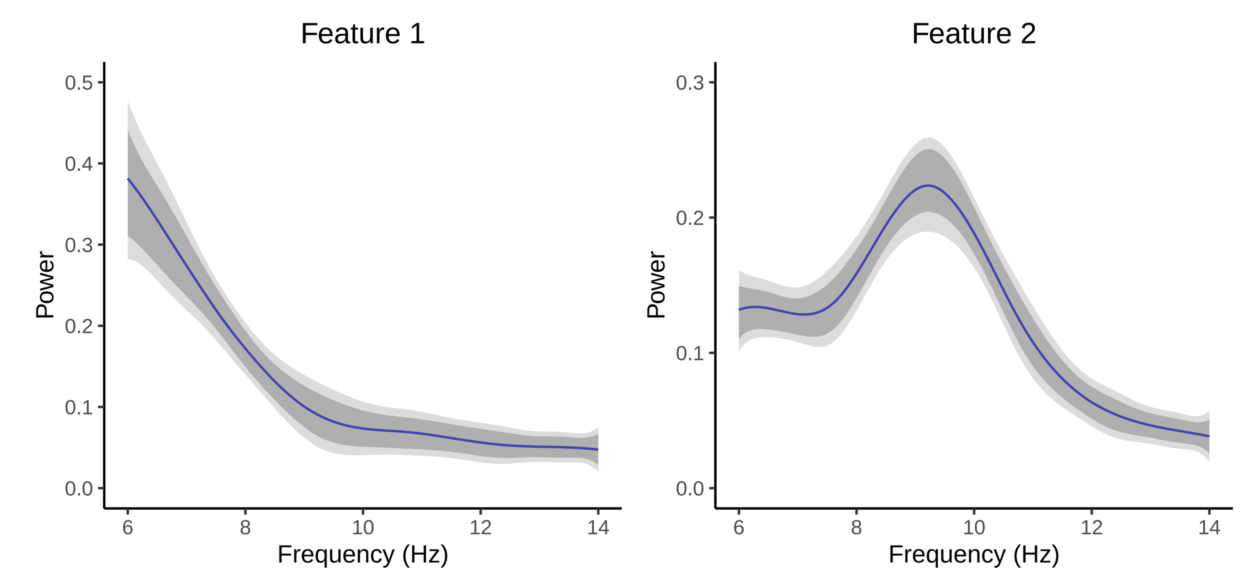


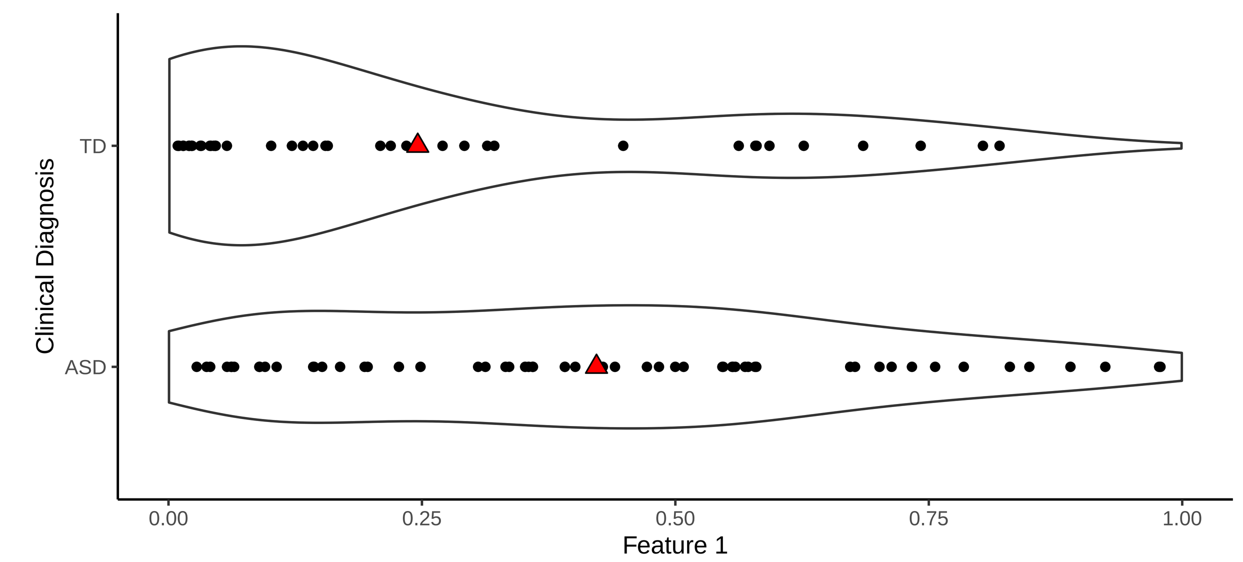
Fig. 3: Observed data from the T8 electrode.

Analysis of the T8 Electrode

In this case study, we fit a two functional feature partial membership model on data from the T8 electrode. The second recovered functional feature can be interpreted as a mature alpha peak, which has been linked to neural development in TD children [4]. The first recovered functional feature is what is known as a $1/f$ noise, or pink noise. $1/f$ noise is expected to be present for all individuals to some degree, however we can see that there is no discernible alpha peak.



We can see that TD children tend to load heavily on the second functional feature, whereas children with ASD have a higher level of heterogeneity.



Remarks

In this project, we proposed a scalable partial membership model that maintains data-driven learning of the covariance structure. Compared to previous work on partial membership models, our proposal allows for increased modeling flexibility, with the benefit of a directly interpretable mean and covariance structure.

References

[1] Abigail Dickinson et al. "Peak alpha frequency is a neural marker of cognitive function across the autism spectrum". In: *European Journal of Neuroscience* 47.6 (2018), pp. 643–651.

[2] Clara Happ and Sonja Greven. "Multivariate functional principal component analysis for data observed on different (dimensional) domains". In: *Journal of the American Statistical Association* 113.522 (2018), pp. 649–659.

[3] Catherine Lord et al. "Autism spectrum disorder". In: *The Lancet* 392.10146 (2018), pp. 508–520.

[4] El Rodrıguez-Martınez et al. "Frequency shift in topography of spontaneous brain rhythms from childhood to adulthood". In: *Cognitive neurodynamics* 11.1 (2017), pp. 23–33.