Predicting Multiple Sclerosis with Functional Spatial Dynamics Data

Functional Derivatives in R

Andrew Belz & Tyler Reese

Department of Statistics
Brigham Young University

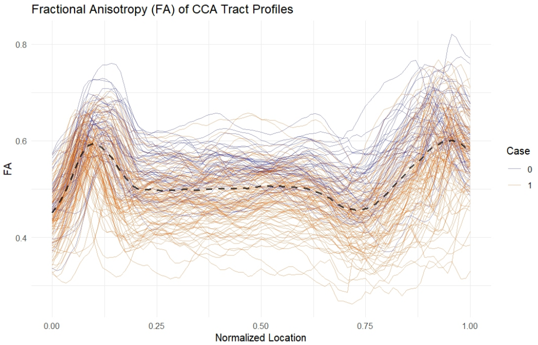
April 22, 2026

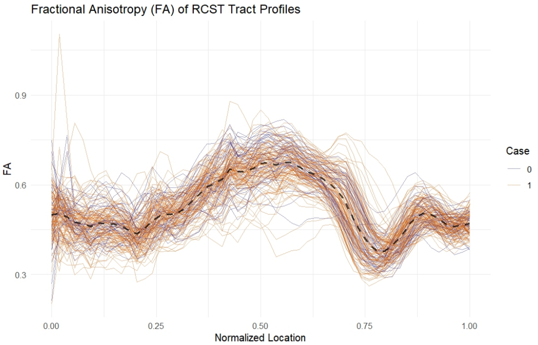
The DTI Dataset

Structure
'data.frame': 382 obs. of 9 variables:
$ ID : num 1001 1002 1003 1004 1005 ...
$ visit : int 1 1 1 1 1 1 1 1 1 1 ...
$ visit.time: int 0 0 0 0 0 0 0 0 0 0 ...
$ Nscans : int 1 1 1 1 1 1 1 1 1 1 ...
$ case : num 0 0 0 0 0 0 0 0 0 0 ...
$ sex : Factor w/ 2 levels "male","female": 2 2 1 1 1 1 1 1 1 1 ...
$ pasat : int NA NA NA NA NA NA NA NA NA NA ...
$ cca : num [1:382, 1:93] 0.491 0.472 0.502 0.402 0.402 ...
..- attr(*, "dimnames")=List of 2
$ rcst : num [1:382, 1:55] 0.257 NaN NaN 0.508 NaN ...
..- attr(*, "dimnames")=List of 2

Table: The DTI Dataset Structure (refund package)

CCA and RCST Tract Profiles





Fractional Anisotropy in Diffusion Tensor Imaging

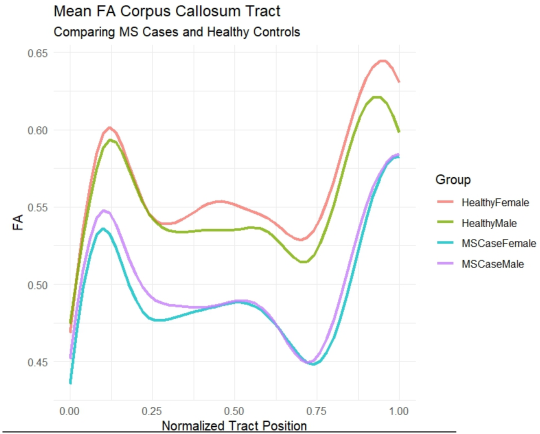
- Fractional Anisotropy (FA) is a numeric measure of how directional water diffuses, in this case specifically in the brain
- High FA - water is well constrained to diffusion along fibers
- Low FA - water diffuses equally in all directions (More common in people with MS)

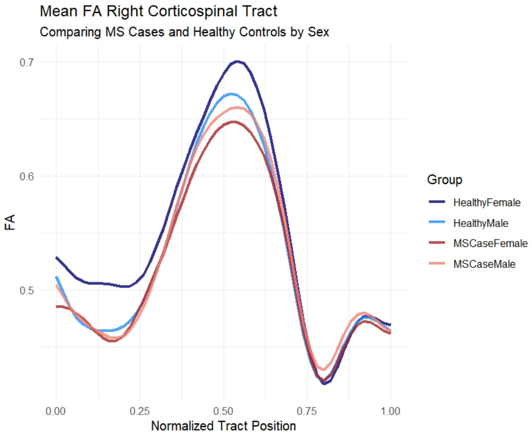
Why is this Relevant to Understanding MS?

- Each curve in our data is several FA measures along a specific tract (CCA or RCST)
- By analyzing the curves, we can predict whether or not someone has MS

Other Features: Sex

- Evidence suggests that mean diffusion differs between men and women[1]

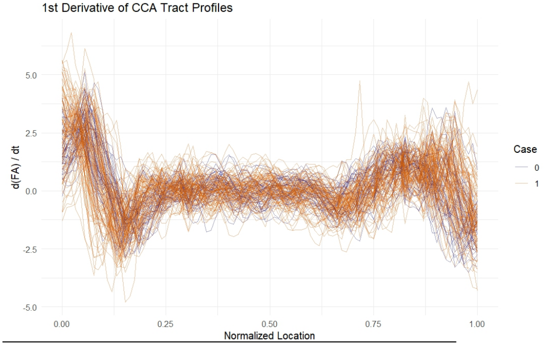


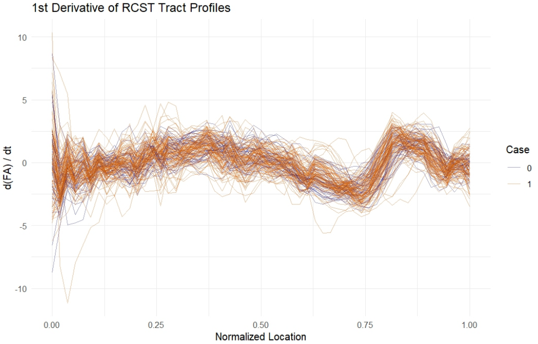


[1] Klistorner et al. 2017.

Why use Derivatives?

- Neurodegeneration occurs non-uniformly[2]
- Highlights relative local differences
- Isolates lesion locations





[2] Meijboom et al. 2023.

Estimating Derivatives from Functional Data

The Karhunen-Loève expansion of X_i:

$$X_i(t) = \mu(t) + \sum_{k=1}^{\infty} \xi_{ik}\phi_k(t)$$

Estimated by

(a) $\hat{\mu}(t)$: smooth mean estimate.

(b) $\hat{\phi}_k(t), \hat{\xi}_{ik}$: estimated via the eigendecomposition of $\hat{G}(s,t)$, a smooth estimate of the covariance operator.

Derivatives of Eigenfunctions (Liu & Miller, 2009)

If X_i is differentiable, then so is its expansion:

$$X_i^{(\nu)}(t) = \mu^{(\nu)}(t) + \sum_{k=1}^{\infty} \xi_{ik} \phi_k^{(\nu)}(t).$$

Derivatives of Eigenfunctions (Liu & Miller, 2009)

If X_i is differentiable, then so is its expansion:

$$X_i^{(\nu)}(t) = \mu^{(\nu)}(t) + \sum_{k=1}^{\infty} \xi_{ik} \phi_k^{(\nu)}(t).$$

Estimate X_i, **then differentiate.**

Eigenfunctions of Derivatives (Dai et al., 2017)

Alternatively, estimate the eigenfunctions of the ν^{th} derivative curve directly:

$$X_i^{(\nu)}(t) = \mu^{(\nu)}(t) + \sum_{k=1}^{\infty} \xi_{ik,\nu} \phi_{k,\nu}(t).$$

Eigenfunctions of Derivatives (Dai et al., 2017)

Alternatively, estimate the eigenfunctions of the ν^{th} derivative curve directly[3]:

$$X_i^{(\nu)}(t) = \mu^{(\nu)}(t) + \sum_{k=1}^{\infty} \xi_{ik,\nu}\phi_{k,\nu}(t).$$

Estimate $X_i^{(\nu)}$, then find the KL-expansion.

[3] Dai, Müller, and Tao 2017.

Which is Better?

$$\sum_{k=1}^{K} \lambda_{k,\nu} \geq \sum_{k=1}^{K} \lambda_k \|\phi_k^{(\nu)}\|_2^2$$

DPCA captures more variance than FPCA with K terms.

Which is Better?

Simulations show that DPCA yields

- better approximation (RMISE)
- same FVE with fewer principal components

Table 1. (Sparse)

$\sigma = 0.5$	$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 5$	FVE	$\hat{\mu}^{(1)}$
FPCA	0.59	0.53	0.44	0.43	0.44	0.44 (4.6)	0.59
DPCA	0.50	0.44	0.43	0.43	0.43	0.43 (2.2)	

$\sigma = 1$	$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 5$	FVE	$\hat{\mu}^{(1)}$
FPCA	0.60	0.54	0.46	0.46	0.46	0.46 (4.5)	0.59
DPCA	0.52	0.47	0.46	0.46	0.46	0.46 (2.2)	

Table 2. (Dense)

$\sigma = 1$	$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 5$	FVE	LOCAL	SMOOTH-DQ
FPCA	0.51	0.42	0.27	0.2	0.16	0.16 (5.0)	0.23	0.65
DPCA	0.32	0.22	0.16	0.13	**0.08**	0.13 (3.9)		

$\sigma = 2$	$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 5$	FVE	LOCAL	SMOOTH-DQ
FPCA	0.51	0.43	0.29	0.27	0.26	0.26 (5.0)	0.51	0.76
DPCA	0.34	0.26	0.22	0.19	**0.18**	0.20 (3.9)		

Fitting DPCA

- FPCAder from fdapace
- kernel smoothing

```
# Take first derivatives of the smoothed curves.
DPCA_cca  <- FPCAder(fpca_fit_cca,  list(p = 1, method = "DPC"))
DPCA_rcst <- FPCAder(fpca_fit_rcst, list(p = 1, method = "DPC"))
```

Figure: fdapace::FPCAder

Functional Logistic Regression: Full Model

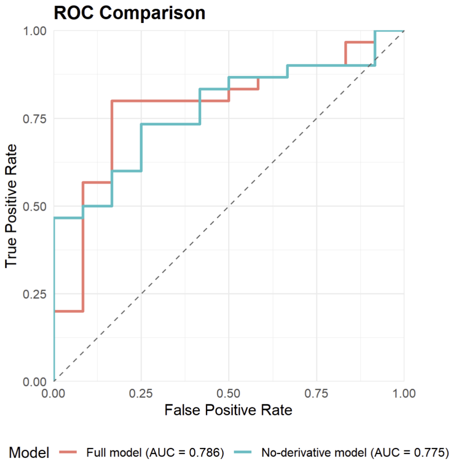
$$\text{logit}\{\overbrace{P(Y_i = 1)}^{\text{positive for MS}}\} = \beta_0 + \sum_{k=1}^{4} \int \overbrace{X_{ki}(t)}^{\text{Curves}} \beta_k(t)\, dt + \gamma \overbrace{Z_i}^{\text{Sex}} + \varepsilon_i,$$

$$X_{1i}(t) = \text{RCSTcurve},$$
$$X_{2i}(t) = \text{CCA curve},$$
$$X_{3i}(t) = \text{RCST derivative},$$
$$X_{4i}(t) = \text{CCA derivative}.$$

Methods

- Models:
  - Full Model (all 4 curves)
  - Original Curves (RCST & CCA)
- Cross-validation parameters:
  - k: number of basis functions
  - m: which derivative to penalize
  - γ: penalty strength
- Train-Test split:
  - 100 train; 42 test
  - equal proportions of MS cases

Classification Results



Model	AUC
Full Model	0.781
No Derivatives	0.775

Classification Results

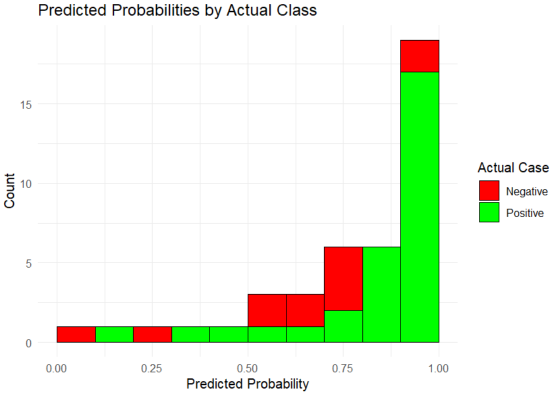
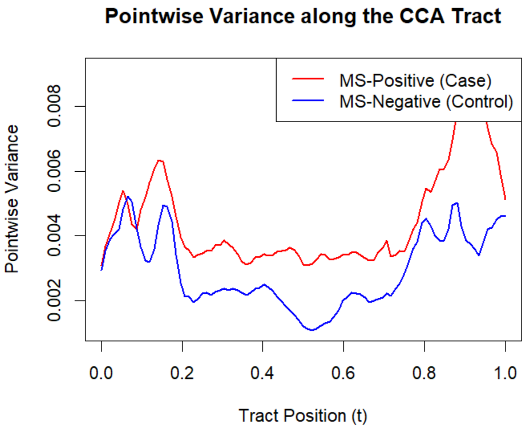


Figure: Enter Caption

Feature Importance

```
Family: binomial
Link function: logit

Formula:
case ~ s(x = cca_smooth.tmat, by = L.cca_smooth, k = 10, bs = "ps",
    m = 2) + s(x = rcst_smooth.tmat, by = L.rcst_smooth, k = 10,
    bs = "ps", m = 2) + s(x = deriv_curves_cca.tmat, by = L.deriv_curves_cca,
    k = 10, bs = "ps", m = 2) + s(x = deriv_curves_rcst.tmat,
    by = L.deriv_curves_rcst, k = 10, bs = "ps", m = 2) + sex

Parametric coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  11.3375     4.8302   2.347   0.0189 *
sexfemale     0.7451     0.7010   1.063   0.2879
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:
                                                  edf Ref.df Chi.sq  p-value
s(cca_smooth.tmat):L.cca_smooth             1.7275880     10 24.006 6.77e-07
s(rcst_smooth.tmat):L.rcst_smooth           1.6368798      8  8.461  0.00407
s(deriv_curves_cca.tmat):L.deriv_curves_cca   0.0003184     10  0.000  0.34481
s(deriv_curves_rcst.tmat):L.deriv_curves_rcst 1.1273177      9  2.791  0.08193

s(cca_smooth.tmat):L.cca_smooth             ***
s(rcst_smooth.tmat):L.rcst_smooth           **
s(deriv_curves_cca.tmat):L.deriv_curves_cca
s(deriv_curves_rcst.tmat):L.deriv_curves_rcst .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) =  0.481   Deviance explained = 45.6%
-REML = 40.802  Scale est. = 1         n = 100
```

Variance


Pointwise Variance along the CCA Tract
MS-Positive (Case)
MS-Negative (Control)
Pointwise Variance
0.002
0.004
0.006
0.008
0.0
0.2
0.4
0.6
0.8
1.0
Tract Position (t)

Thank you!

References I

Dai, Xiongtao, Hans-Georg Müller, and Wenwen Tao (2017). *Derivative Principal Component Analysis for Representing the Time Dynamics of Longitudinal and Functional Data*. arXiv: 1707.04360 [stat.ME]. URL: https://arxiv.org/abs/1707.04360.

Klistorner, A. et al. (2017). "Evidence of progressive tissue loss in the core of chronic MS lesions: A longitudinal DTI study". In: *NeuroImage: Clinical* 17, pp. 1028–1035. DOI: 10.1016/j.nicl.2017.12.010.

Meijboom, R. et al. (2023). "Patterns of brain atrophy in recently-diagnosed relapsing-remitting multiple sclerosis". In: *PloS one* 18.7, e0288967. DOI: 10.1371/journal.pone.0288967.