Generalized Linear Model for a Spherical Response with Projection-based Inference

Kipoong Kim[1] Jörn Schulz[2] Sungkyu Jung[3]

[1]Department of Statistics, Changwon National University

[2]Department of Mathematics and Physics, University of Stavanger

[3]Department of Statistics, Seoul National University

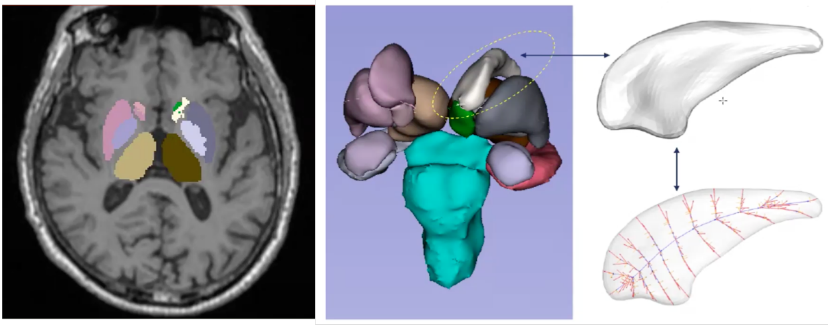
August 21, 2025

Table of Contents

Motivating data

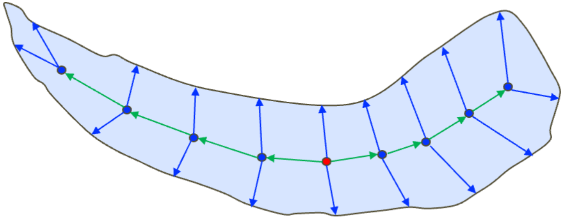
Shape data from a brain MRI

- From Brain MRI, we can obtain the shape of specific regions as 3D point cloud data in (x,y,z) coordinates. In our case, the shapes of hippocampus were collected.
- In this study, we aim test for shape differences in the hippocampus, derived from a brain MRI, between a Parkinson's disease (PD) group and a normal control group.



Data Representation

- However, 3d point cloud data originally poses significant analytical challenges, as the large number of points effectively creates a high-dimensional feature space.
- To address this, we transform the shape into a compressed representation using the locally parametrized discrete-skeletal representation (LP-ds-rep) method (Taheri and Schulz, 2021).
- Simply speaking, this transformation gives us key features that effectively capture the shape: the centroid, connection vectors forming the spine, and spoke vectors extending toward the shape's boundary.



Mathematical Formulation

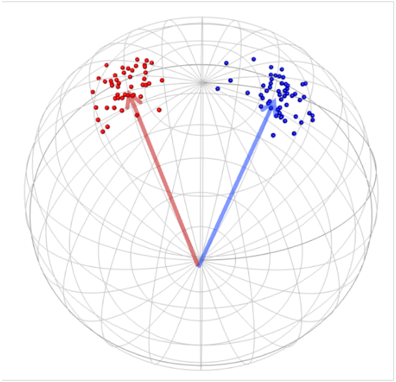
- Connection and spoke vectors can be thought of as a direction times its length.
- Each data point can be considered an element of the following space:

$$\underbrace{\mathbb{R}^3}_{\text{centroid}} \times \underbrace{(\mathbb{S}^2)^{n_c} \times (\mathbb{R}_+)^{n_c}}_{\text{connections}} \times \underbrace{(\mathbb{S}^2)^{n_s} \times (\mathbb{R}_+)^{n_s}}_{\text{spokes}}, \tag{1}$$

  where $\mathbb{S}^{q-1} = \left\{\boldsymbol{y} \in \mathbb{R}^q : \boldsymbol{y}^T\boldsymbol{y} = 1\right\}$.
- In this work, we focus specifically on the spoke vectors. We propose a method to test for differences in a single spoke direction between the case and control groups.

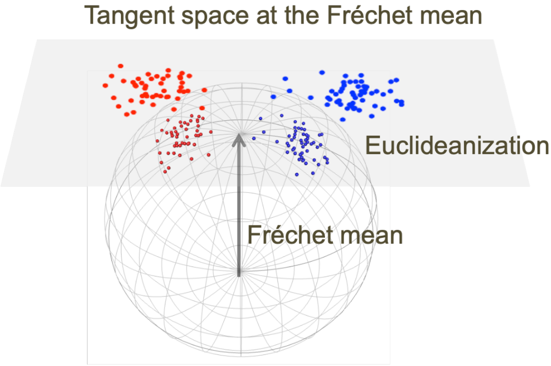
The Two-Sample Problem on a Sphere

- Since each spoke direction is a unit vector, it lies on $\mathbb{S}^2$. The problem of testing for a difference in spoke directions between the two groups can be visualized as follows:



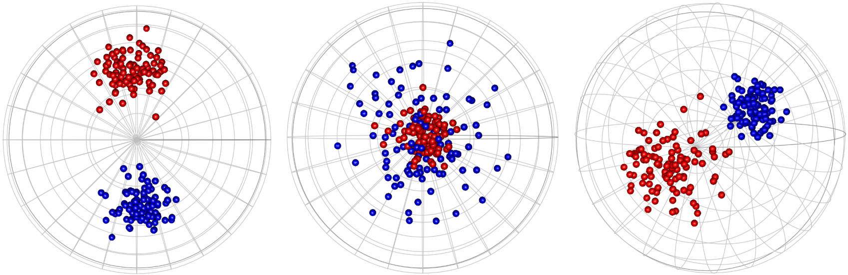
An approach to test the difference

- (Tangent space) Calculate the Fréchet mean and consider the Tangent space at the Fréchet mean
- (Euclideanization) Map the data points to the tangent space at the Fréchet mean (via the logarithmic map).
- (Testing) Apply a standard Euclidean testing method, such as Hotelling's T^2-test.



Research question

Location and/or dispersion problem



(a) Location (b) Dispersion (c) Both

- Existing regression models for a spherical response
  - Fisher & Lee (1992); Presnell et al. (1998): Limited to $\mathbb{S}^1$
  - Paine et al. (2020): No inferential procedure (potentially available)

Methods

von Mises-Fisher (vMF) distribution

- For $\boldsymbol{y} \in \mathbb{S}^{q-1}$, the probability density function of von Mises-Fisher distribution (vMF) is defined as

$$f_{vMF}(\boldsymbol{y} : \boldsymbol{\zeta}, \kappa) = C_q(\kappa) \exp\left(\kappa \boldsymbol{\zeta}^T \boldsymbol{y}\right)$$

  for a concentration parameter $\kappa \geq 0$ and mean direction $\boldsymbol{\zeta} \in \mathbb{S}^{q-1}$.
  - Here, $C_q(\kappa)$ is a normalization constant defined as

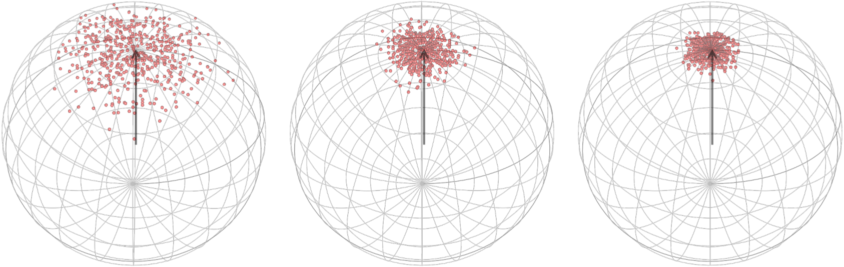
$$C_q(\kappa) = \frac{\kappa^{q/2-1}}{(2\pi)^{q/2} I_{q/2-1}(\kappa)},$$

    where $I_\nu(\cdot)$ is the modified Bessel function of the first kind at order ν.

von Mises-Fisher (vMF) distribution

- An example of von Mises-Fisher distribution with

$$q = 3, \quad \boldsymbol{\zeta} = (0, 0, 1)^\top, \quad \kappa \in \{5, 10, 15\}$$



(a) $\kappa = 5$ (b) $\kappa = 10$ (c) $\kappa = 15$

von Mises-Fisher (vMF) distribution

- Constraint on a mean direction parameter
  - A mean direction $\boldsymbol{\zeta}$ of vMF belongs to $\mathbb{S}^{q-1}$, so it is not directly linked to a linear predictor $\boldsymbol{\eta}_i = \boldsymbol{\mu} + \sum_{j=1}^{p} x_{ij}\boldsymbol{\beta}_j \in \mathbb{R}^q$

$$\boldsymbol{\zeta}_i \overset{?}{=} \boldsymbol{\mu} + \sum_{j=1}^{p} x_{ij}\boldsymbol{\beta}_j$$

    for $\boldsymbol{\mu} \in \mathbb{R}^q$.
  - A concentration parameter κ of vMF is also constrained

$$\kappa \in \mathbb{R}_+ := \{c \in \mathbb{R} : c \geq 0\}$$

- We want to link the mean direction to a linear predictor without constraints.

A double regression model for location and dispersion

- For a linear predictor $\boldsymbol{\eta}_i$, we can consider two types of link functions:
  - Location link:

$$\boldsymbol{\zeta}_i = \frac{\boldsymbol{\eta}_i^{\mathsf{loc}}}{\|\boldsymbol{\eta}_i^{\mathsf{loc}}\|_2} \quad \Rightarrow \quad g_1(\boldsymbol{\zeta}_i) = \boldsymbol{\eta}_i^{\mathsf{loc}} = \boldsymbol{\mu}^{\mathsf{loc}} + \sum_{j=1}^{p} x_{ij}\boldsymbol{\beta}_j^{\mathsf{loc}}$$

  - Dispersion link:

$$\kappa_i = \|\boldsymbol{\eta}_i^{\mathsf{disp}}\| \quad \Rightarrow \quad g_2(\kappa_i) = \boldsymbol{\eta}_i^{\mathsf{disp}} = \boldsymbol{\mu}^{\mathsf{disp}} + \sum_{j=1}^{p} x_{ij}\boldsymbol{\beta}_j^{\mathsf{disp}}$$

- The log-likelihood can be written as

$$\begin{aligned}\ell(\boldsymbol{\mu}^{\mathsf{loc}}, \boldsymbol{\mu}^{\mathsf{disp}}, \mathbf{B}^{\mathsf{loc}}, \mathbf{B}^{\mathsf{disp}} | \mathbf{Y}) &= \sum_{i=1}^{n} \kappa_i \boldsymbol{\zeta}_i^{\top} \boldsymbol{y}_i + \log C_q(\kappa_i) \\ &= \sum_{i=1}^{n} \|\boldsymbol{\eta}_i^{\mathsf{disp}}\|_2 \frac{{\boldsymbol{\eta}_i^{\mathsf{loc}}}^{\top} \boldsymbol{y}_i}{\|\boldsymbol{\eta}_i^{\mathsf{loc}}\|_2} + \log C_q(\|\boldsymbol{\eta}_i^{\mathsf{disp}}\|_2)\end{aligned} \tag{2}$$

A shared parameter model for location and dispersion

- Alternatively, we can consider the model with shared parameter $\boldsymbol{\beta}_j$ between location and dispersion link functions:
  - Location link:

$$\boldsymbol{\zeta}_i = \frac{\boldsymbol{\eta}_i}{\|\boldsymbol{\eta}_i\|_2} \quad \Rightarrow \quad g_1(\boldsymbol{\zeta}_i) = \boldsymbol{\eta}_i = \boldsymbol{\mu} + \sum_{j=1}^{p} x_{ij}\boldsymbol{\beta}_j$$

  - Dispersion link:

$$\kappa_i = \|\boldsymbol{\eta}_i\| \quad \Rightarrow \quad g_2(\kappa_i) = \boldsymbol{\eta}_i = \boldsymbol{\mu} + \sum_{j=1}^{p} x_{ij}\boldsymbol{\beta}_j$$

- The log-likelihood can be written as

$$\begin{aligned} \ell(\boldsymbol{\mu}, \mathbf{B}|\mathbf{Y}) &= \sum_{i=1}^{n} \kappa_i \boldsymbol{\zeta}_i^\top \boldsymbol{y}_i + \log C_q(\kappa_i) \\ &= \sum_{i=1}^{n} \boldsymbol{\eta}_i{}^\top \boldsymbol{y}_i + \log C_q(\|\boldsymbol{\eta}_i\|_2) \end{aligned} \quad (3)$$

Reparameterization

- The shared parameter model is equivalent to the model with reparameterization $\boldsymbol{\theta} = \kappa \cdot \boldsymbol{\zeta} \in \mathbb{R}^q$
- This leads to a reparameterized vMF distribution with a mean vector $\boldsymbol{\theta} \in \mathbb{R}^q$:
$$f_{\mathrm{vMF}}(\boldsymbol{y} : \boldsymbol{\theta}) = \exp\left(\boldsymbol{\theta}^T \boldsymbol{y} + \log C_q(\|\boldsymbol{\theta}\|)\right). \tag{4}$$
- This reparameterization reveals that the vMF distribution is a member of the canonical exponential family with the natural parameter $\boldsymbol{\theta}$.
- More specifically, by comparing the canonical exponential family form
$$f(\boldsymbol{y}|\boldsymbol{\theta}, \phi) = \exp\left(\frac{\boldsymbol{\theta}^T \boldsymbol{y} - b(\boldsymbol{\theta})}{a(\phi)} + c(\boldsymbol{y}, \phi)\right),$$
it turns out that
$$a(\phi) = 1,\ b(\boldsymbol{\theta}) = -\log C_q(\|\boldsymbol{\theta}\|),\ \text{and}\ c(\boldsymbol{y}, \phi) = 0$$

Proposition 1

For $\boldsymbol{\theta} \in \mathbb{R}^q$ *with* $\|\boldsymbol{\theta}\| \neq 0$, *the first and second derivatives of* $b(\boldsymbol{\theta})$ *can be expressed as*

$$b'(\boldsymbol{\theta}) = \frac{\boldsymbol{\theta}}{\|\boldsymbol{\theta}\|} \left\{ \frac{I_{q/2}(\|\boldsymbol{\theta}\|)}{I_{q/2-1}(\|\boldsymbol{\theta}\|)} \right\}$$

$$b''(\boldsymbol{\theta}) = \left(\frac{\boldsymbol{\theta}}{\|\boldsymbol{\theta}\|} \right) \left(\frac{\boldsymbol{\theta}}{\|\boldsymbol{\theta}\|} \right)^T H_q(\|\boldsymbol{\theta}\|) + \frac{B_q(\|\boldsymbol{\theta}\|)}{\|\boldsymbol{\theta}\|} \left\{ \mathbf{I} - \left(\frac{\boldsymbol{\theta}}{\|\boldsymbol{\theta}\|} \right) \left(\frac{\boldsymbol{\theta}}{\|\boldsymbol{\theta}\|} \right)^T \right\},$$

where $B_q(\kappa) = \dfrac{I_{q/2}(\kappa)}{I_{q/2-1}(\kappa)}$ *is the ratio of the modified Bessel functions (also known as the Mean Resultant Length) and*

$$H_q(\kappa) = 1 - B_q^2(\kappa) - \frac{q-1}{\kappa} B_q(\kappa)$$

Linking the spherical responses to covariates

- We can consider the generalized linear model for the reparametrized vMF distribution as

$$g\left(\mathbb{E}[\boldsymbol{y}_i|\boldsymbol{x}_i]\right) = \boldsymbol{\theta}_i = \boldsymbol{\mu} + \sum_{j=1}^{p} x_{ij}\boldsymbol{\beta}_j = \boldsymbol{\mu} + \mathbf{B}^T\boldsymbol{x}_i$$

$$\boldsymbol{y}_i|\boldsymbol{x}_i \sim \mathrm{vMF}(\boldsymbol{\theta}_i),$$

where

  - $\boldsymbol{y}_i \in \mathbb{S}^{q-1}, \quad \boldsymbol{\mu}, \boldsymbol{\theta}_i \in \mathbb{R}^q$,
  - $g(\boldsymbol{\xi}) = B_q^{-1}(\|\boldsymbol{\xi}\|)\frac{\boldsymbol{\xi}}{\|\boldsymbol{\xi}\|}$ is a known link function,
  - $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip})^T$ is a p-dimensional covariate vector, and $\mathbf{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)^T$ is a column-centered design matrix.
  - $\mathbf{B} = [\mathbf{B}_1, \ldots, \mathbf{B}_q] \in \mathbb{R}^{p\times q}$ is a a coefficient matrix whose the (j,k)-th element represents the effect of j-th covariate on the k-th response.

- Let us define
  - $\tilde{\boldsymbol{x}}_i := \boldsymbol{x}_i \otimes \mathbf{I}_q \in \mathbb{R}^{pq \times q}$
  - $\boldsymbol{\beta}^* := \mathrm{vec}(\mathbf{B}^T) = (\boldsymbol{\beta}_1^T, \boldsymbol{\beta}_2^T, \ldots, \boldsymbol{\beta}_p^T)^T \in \mathbb{R}^{pq \times 1}$
- Then, we can write the model as

$$\boldsymbol{\theta}_i = \boldsymbol{\mu} + \tilde{\boldsymbol{x}}_i^T \boldsymbol{\beta}^*. \tag{5}$$

and its log-likelihood as

$$\ell(\boldsymbol{\mu}, \boldsymbol{\beta}^*) = \sum_{i=1}^{n} \left\{ \boldsymbol{\mu}^T \boldsymbol{y}_i + \boldsymbol{\beta}^{*T} \tilde{\boldsymbol{x}}_i \boldsymbol{y}_i - b(\boldsymbol{\mu} + \tilde{\boldsymbol{x}}_i^T \boldsymbol{\beta}^*) \right\}, \tag{6}$$

- Let us denote the full parameter vector by $\boldsymbol{\gamma} = (\boldsymbol{\mu}^\top, \boldsymbol{\beta}^{*\top})^\top$.
- The corresponding score function $\boldsymbol{s}_n(\boldsymbol{\gamma})$ and Fisher information matrix $\mathbf{F}_n(\boldsymbol{\gamma})$ are given by

Proposition 2

$$\boldsymbol{s}_n(\boldsymbol{\gamma}) = \frac{\partial \ell(\boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}} = \sum_{i=1}^{n} \tilde{\boldsymbol{x}}_i \left\{ \boldsymbol{y}_i - b'(\boldsymbol{\mu} + \tilde{\boldsymbol{x}}_i^T \boldsymbol{\beta}^*) \right\}$$

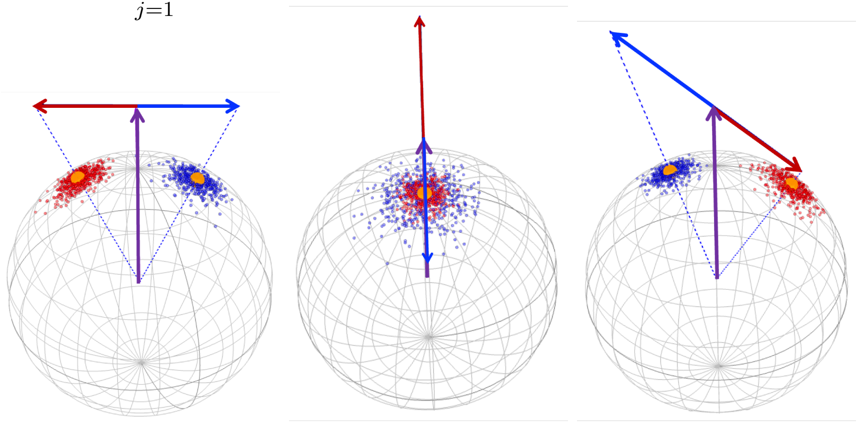
$$\mathbf{F}_n(\boldsymbol{\gamma}) = \frac{\partial^2 \ell(\boldsymbol{\gamma})}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}^T} = \sum_{i=1}^{n} \left\{ \tilde{\boldsymbol{x}}_i b''(\boldsymbol{\mu} + \tilde{\boldsymbol{x}}_i^T \boldsymbol{\beta}^*) \tilde{\boldsymbol{x}}_i^T \right\}.$$

Proposition 3

If $\sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i^T$ is of full rank, the Fisher information matrix is positive definite, and the MLE of $\boldsymbol{\gamma}$ is unique

- Recall that

$$\boldsymbol{\theta}_i = \boldsymbol{\mu} + \sum_{j=1}^{p} x_{ij}\boldsymbol{\beta}_j \in \mathbb{R}^q, \qquad \boldsymbol{y}_i|\boldsymbol{x}_i \sim \mathrm{vMF}(\boldsymbol{\theta}_i) \text{ with } x_{ij} \in \{-1,+1\}\ ,$$



(a) orthogonal (b) parallel (c) oblique

Maximum likelihood estimation with or without constraints

- Unconstrained Maximum likelihood estimation (MLE):

$$\underset{\boldsymbol{\mu}\in\mathbb{R}^q,\ \boldsymbol{\beta}^*\in\mathbb{R}^{pq}}{\arg\max} \sum_{i=1}^{n} \left\{ (\boldsymbol{\mu} + \tilde{\boldsymbol{x}}_i^T \boldsymbol{\beta}^*)^T \boldsymbol{y}_i + \log C_q(\|\boldsymbol{\mu} + \tilde{\boldsymbol{x}}_i^T \boldsymbol{\beta}^*\|) \right\}$$

- Constrained MLE under orthogonality constraints:

$$\underset{\boldsymbol{\mu}\in\mathbb{R}^q,\ \boldsymbol{\beta}^*\in\mathbb{R}^{pq}}{\arg\max} \sum_{i=1}^{n} \left\{ (\boldsymbol{\mu} + \tilde{\boldsymbol{x}}_i^T \boldsymbol{\beta}^*)^T \boldsymbol{y}_i + \log C_q(\|\boldsymbol{\mu} + \tilde{\boldsymbol{x}}_i^T \boldsymbol{\beta}^*\|) \right\}$$

$$\text{subject to } \boldsymbol{\beta}_j \perp \boldsymbol{\mu} \ \ \forall j = 1, \ldots, p$$

  - If the true parameters satisfy this constraint, the constrained estimator can be more efficient (i.e., have lower variance).

- Let us denote the constrained estimator of $\boldsymbol{\gamma}$ by $\hat{\boldsymbol{\gamma}}^c$ and the unconstrained version by $\hat{\boldsymbol{\gamma}}^{uc}$. Then, we have

Theorem 1

☐ *Under the regularity conditions, the weak consistency and the asymptotic normality of* $\hat{\boldsymbol{\gamma}}^c$ *can be obtained as follows:*

$$\hat{\boldsymbol{\gamma}}^{uc} \xrightarrow{p} \boldsymbol{\gamma}_0 \quad \textit{and} \quad \hat{\boldsymbol{\gamma}}^{uc} \xrightarrow{d} N(\boldsymbol{\gamma}_0, \mathbf{F}(\boldsymbol{\gamma}_0)^{-1}).$$

☐ *By further assuming that* $\boldsymbol{\beta}_{j,0} \perp \boldsymbol{\mu}_0 \; \forall j$*, we then have*

$$\hat{\boldsymbol{\gamma}}^{c} \xrightarrow{p} \boldsymbol{\gamma}_0 \quad \textit{and} \quad \hat{\boldsymbol{\gamma}}^{c} \xrightarrow{d} N(\boldsymbol{\gamma}_0, \boldsymbol{\Sigma}_c),$$

where the asymptotic covariance matrix $\boldsymbol{\Sigma}_c$ *is given by*

$$\mathbf{F}(\boldsymbol{\gamma}_0)^{-1} - \mathbf{F}(\boldsymbol{\gamma}_0)^{-1}\mathbf{H}_0^\top \left(\mathbf{H}_0\mathbf{F}(\boldsymbol{\gamma}_0)^{-1}\mathbf{H}_0^\top\right)^{-1} \mathbf{H}_0\mathbf{F}(\boldsymbol{\gamma}_0)^{-1}$$

and $\mathbf{H}_0 = \mathbf{H}_0(\boldsymbol{\gamma}_0)$ *is the Jacobian of the constraints w.r.t.* $\boldsymbol{\gamma}$ *at* $\boldsymbol{\gamma}_0$*.*

Corollary 2 (Asymptotic Normality via Projection)

Let us let $\mathbf{P}_{\hat{\boldsymbol{\mu}}} = \hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^T/\|\hat{\boldsymbol{\mu}}\|^2,\ \mathbf{P}_{\hat{\boldsymbol{\mu}}^\perp} = \mathbf{I} - \mathbf{P}_{\hat{\boldsymbol{\mu}}}$ *be the projection matrix onto* $\hat{\boldsymbol{\mu}}$ *and its orthogonal complement, respectively.*

☐ *Under the regularity conditions, as* $n \to \infty$*, we have*

$$\mathbf{P}_{\hat{\boldsymbol{\mu}}^\perp}\hat{\boldsymbol{\beta}}_j^{uc}|\hat{\boldsymbol{\mu}} \xrightarrow{d} N\left(\mathbf{P}_{\boldsymbol{\mu}^\perp}\boldsymbol{\beta}_j^{uc},\ \boldsymbol{\Sigma}_{\boldsymbol{\mu}^\perp}\right)$$

where $\boldsymbol{\Sigma}_{\boldsymbol{\mu}^\perp} = \mathbf{P}_{\boldsymbol{\mu}^\perp}\boldsymbol{c}_j^T\left[\mathbf{F}_{0,22}^{-1} - \mathbf{F}_{0,21}^{-1}\mathbf{F}_{0,11}\mathbf{F}_{0,12}^{-1}\right]\boldsymbol{c}_j\mathbf{P}_{\boldsymbol{\mu}^\perp}^T$ *and* $\boldsymbol{c}_j$ *is the index matrix extracting the components representing the* j*-th covariate.*

☐ *Similarly, for the* μ*-directional projection, we have the following result:*

$$\mathbf{P}_{\hat{\boldsymbol{\mu}}}\hat{\boldsymbol{\beta}}_j^{uc}|\hat{\boldsymbol{\mu}} \xrightarrow{d} N\left(\mathbf{P}_{\boldsymbol{\mu}}\boldsymbol{\beta}_j,\ \boldsymbol{\Sigma}_{\boldsymbol{\mu}}\right)$$

where $\boldsymbol{\Sigma}_{\boldsymbol{\mu}} = \mathbf{P}_{\boldsymbol{\mu}}\boldsymbol{c}_j^T\left[\mathbf{F}_{0,22}^{-1} - \mathbf{F}_{0,21}^{-1}\mathbf{F}_{0,11}\mathbf{F}_{0,12}^{-1}\right]\boldsymbol{c}_j\mathbf{P}_{\boldsymbol{\mu}}^T.$

Confidence region

- Confidence regions of the unconstrained MLE, projected version onto $\boldsymbol{\mu}$ and its orthogonal complement.

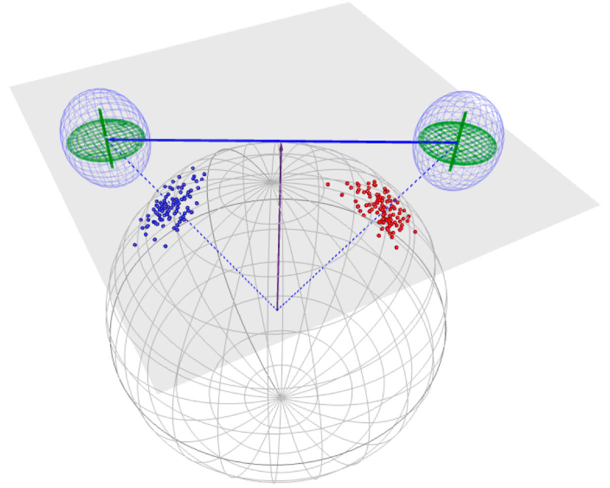


Figure: True coefficient vector $\boldsymbol{\beta}_{0,1}$ is orthogonal to true mean vector $\boldsymbol{\mu}_0$

Simulation studies: two-sample test

- $n \in \{50, 100, 500, 1000\}$
- $q = 3$
- $\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right]$
  - $\rho \in \{0.2, 0.8\}$
- $X_1 = 2 \cdot \mathbb{1}(Z_1 \geq 0) - 1$
- $X_2 = Z_2$
- Data generating model:

$$\boldsymbol{y}_i \sim \mathrm{vMF}(\boldsymbol{\theta}_i),$$

  where $\boldsymbol{\theta}_i = \boldsymbol{\mu} + X_{i1}\boldsymbol{\beta}_1 + X_{i2}\boldsymbol{\beta}_2$ for $i = 1, \ldots, n$.

Simulation studies: two-sample test

- $\boldsymbol{\mu} = (0, 0, 20)^T$

- $\boldsymbol{\beta}_j = \underbrace{\frac{\boldsymbol{d}_j}{\|\boldsymbol{d}_j\|_2}}_{\text{direction}} \times \underbrace{\mathcal{N}(0, s_j^2)}_{\text{length}}$ for $j = 1, 2$
  - Direction: $\boldsymbol{d}_j \in \text{span}(\boldsymbol{\mu})$ or $\boldsymbol{d}_j \in \text{span}(\boldsymbol{\mu})^{\perp}$,
    - $(\boldsymbol{\beta}_1 \in \text{span}(\boldsymbol{\mu})^{\perp}$, $\boldsymbol{\beta}_2 \in \text{span}(\boldsymbol{\mu})^{\perp})$
    - $(\boldsymbol{\beta}_1 \in \text{span}(\boldsymbol{\mu})^{\perp}$, $\boldsymbol{\beta}_2 \in \text{span}(\boldsymbol{\mu}))$
    - $(\boldsymbol{\beta}_1 \in \text{span}(\boldsymbol{\mu})$, $\boldsymbol{\beta}_2 \in \text{span}(\boldsymbol{\mu})^{\perp})$
    - $(\boldsymbol{\beta}_1 \in \text{span}(\boldsymbol{\mu})$, $\boldsymbol{\beta}_2 \in \text{span}(\boldsymbol{\mu}))$
  - Effect size: $s_j \in \{0, 2, 5\}$ for $j = 1, 2$

- The number of simulations = 100

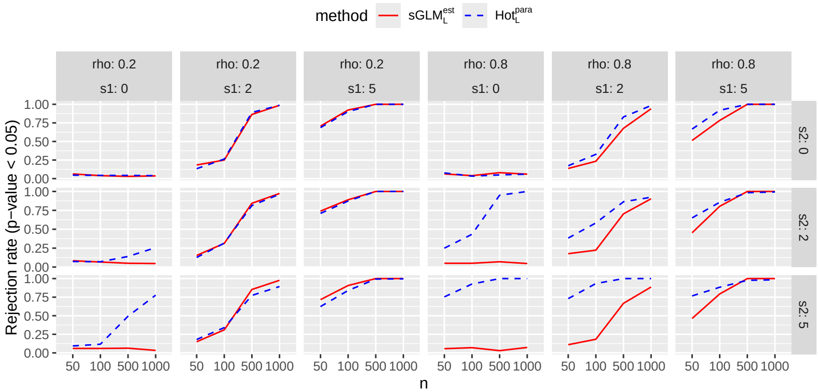
Comparing methods

- $\mathrm{sGLM}_{\mathrm{L}}^{\hat{\boldsymbol{\mu}}^{\perp}}$: Location model of sphereGLM projected onto the orthogonal complement of the estimated $\hat{\boldsymbol{\mu}}$
- $\mathrm{sGLM}_{\mathrm{D}}^{\hat{\boldsymbol{\mu}}}$: Dispersion model of sphereGLM projected onto the estimated $\hat{\boldsymbol{\mu}}$
- $\mathrm{Hot}_{\mathrm{L}}^{\mathrm{para}}$: Hotelling's T^2 test
- $\mathrm{Disp}_{\mathrm{D}}^{\mathrm{para}}$: Distance-based dispersion test[1]

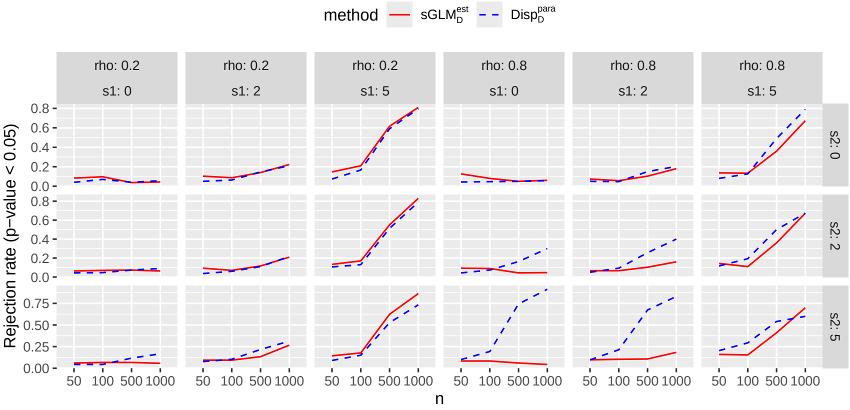
where $\#$ of perms $= 5000$

- Evaluation criteria
  - (i) Type I error rate and (ii) Statistical power

$(\boldsymbol{\beta}_1 \in \operatorname{span}(\boldsymbol{\mu})^{\perp}, \boldsymbol{\beta}_2 \in \operatorname{span}(\boldsymbol{\mu})^{\perp})$

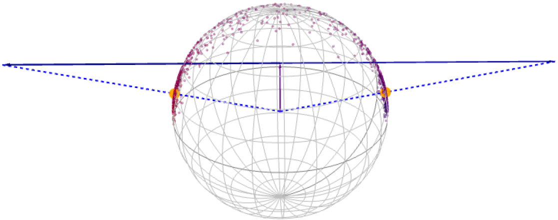


$(\boldsymbol{\beta}_1 \in \mathrm{span}(\boldsymbol{\mu}), \, \boldsymbol{\beta}_2 \in \mathrm{span}(\boldsymbol{\mu}))$



Concluding remarks

- A limitation arises with continuous covariates:
  - The domain of the conditional expectation $\mathbb{E}[\boldsymbol{y}_i | x_i]$ may be restricted to a hemisphere. As a result, it would be difficult to account for spherical response



  - As the parameter $\boldsymbol{\theta}_i$ varies with the covariate, the location and dispersion effects become entangled. Disentangling these two effects remains a task for future work.

- Future work includes extending the model to multivariate responses on product manifolds, such as the torus ($\mathbb{S}^1 \times \mathbb{S}^1$), products of spheres ($\mathbb{S}^2 \times \mathbb{S}^2$), or a cylinder ($\mathbb{S}^2 \times \mathbb{R}_+$).

Thank you for your attention ! ☺