Generalized Fiducial Inference on Riemannian Manifolds

Alexander C Murph[1]; Jan Hannig[1]; Jonathan P Williams[2]
[1]University of NC at Chapel Hill, [2]NC State University

How do you define a fiducial distribution on a manifold $\mathcal{M}$?

Express your data **y** in terms of a known random quantity *u* and unknown parameter vector θ through a **Data Generating Algorithm:** y = A(*u*,θ).

Generalized Fiducial Distribution (GFD) [1]

$$r_{\mathbf{y}}(\theta) = \frac{f(\mathbf{y}|\theta)J(\mathbf{y},\theta)}{\int f(\mathbf{y}|\theta')\,J(\mathbf{y},\theta')\,d\theta'}$$

where

$f(\mathbf{y}|\theta)$ → **the data likelihood**

$J(\mathbf{y},\theta)$ → **change of variables Jacobian**

$$J(\mathbf{y},\theta) = D\left(\nabla_\theta A\left(u,\theta\right)|_{u=A^{-1}(\mathbf{y},\theta)}\right)$$

$D(M) = (\det M'M)^{\frac{1}{2}}$ → **det based on the l2 norm**

Constrained Generalized Fiducial Distribution

$$r_{\mathbf{y},\mathcal{M}}(\theta|\mathbf{y}) = \frac{f(y|\theta)D^\star\left(\nabla_\theta A(u,\theta)|_{u=A^{-1}(\mathbf{y},\theta)}P_\theta\right)}{\int_{\mathcal{M}} f(y|\theta^*)D^\star\left(\nabla_{\theta^*} A(u,\theta^*)|_{u=A^{-1}(\mathbf{y},\theta^*)}P_{\theta^*}\right)d\lambda(\theta^*)}, \qquad \theta\in\mathcal{M},$$

$$P_\theta := I_d - (\nabla_\theta g)'(\nabla_\theta g(\nabla_\theta g)')^{-1}\nabla_\theta g,$$

$D^\star(M) = (\mathrm{pdet}\, M'M)^{\frac{1}{2}}$ → **pseudodeterminant analogue to D(M)**

Assume there exists a differentiable function **g** such that θ ∈ M ↔ g(θ) = 0. When our parameter space is M, the gradient must be updated!

Two Approaches

Our approach against the direct solution with a parameterization.

Constrained GFD

Sample from $r_{\mathbf{y},\mathcal{M}}(\theta|\mathbf{y})$, calculating the projection P_θ at each point on the manifold.

This **only** uses the information from **g**, and does **not require a parameterization of the manifold!**

For this simple example, we can compare to a direct solution (below).

GFD with Parameterization

Sample from the usual GFD on the **parameterized space**. By the chain rule,

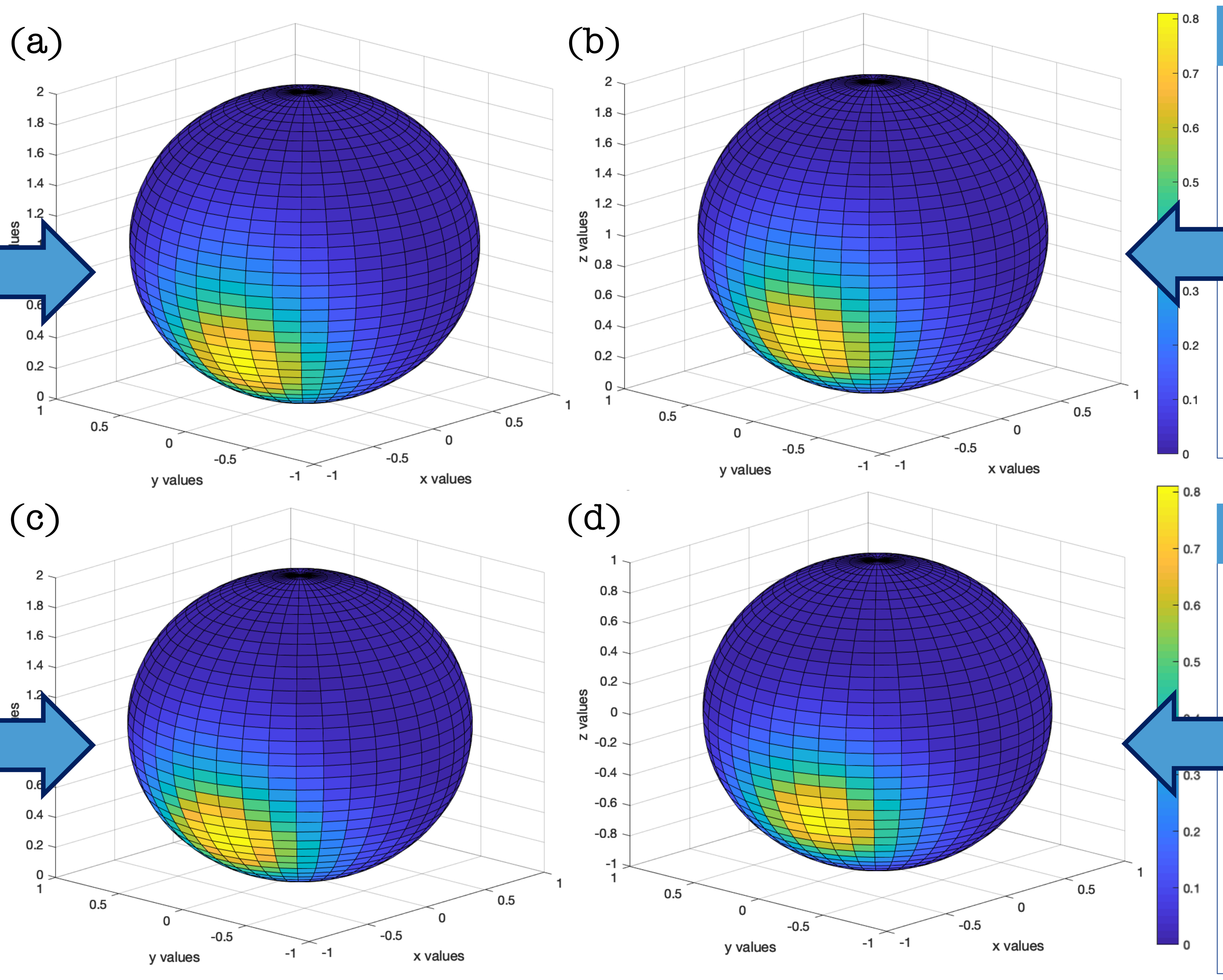
$$r_{\mathbf{y}} \propto f(\Psi(\varphi,\zeta)|\mathbf{y})D(\nabla_\theta A(\mathbf{y},\theta)\nabla_{\varphi,\zeta}\Psi(\varphi,\zeta))$$

where

$$\Psi(\varphi,\zeta) = \left(\cos\varphi\sin\zeta \quad \sin\varphi\sin\zeta \quad \cos\zeta\right)'$$

for $0\le\varphi\le 2\pi,\ 0\le\zeta\le\pi$

This is *not* new theory.



Two MCMC Algorithms

How we sample our Constrained GFD on a constrained space

Constrained Hamiltonian MCMC

Simulates a Hamiltonian's movement through the system, using the RATTLE [2] algorithm to satisfy the constraints

$$0 = g(\theta) \quad \text{and} \quad 0 = (\nabla_\theta g(\theta))\frac{\partial H}{\partial p}$$

For the Hamiltonian [3],

$$H(\theta,p) = \left(\frac{1}{2}p'M(\theta)^{-1}p\right) + \left(\frac{1}{2}\log|M(\theta)| - \log\pi(\theta)\right) + \lambda' g(\theta)$$

$M(\theta)$, user-defined mass matrix, $\pi(\theta) := r_{\mathbf{y},\mathcal{M}}(\theta|\mathbf{y})$

Constrained Metropolis-Hastings

New samples are drawn using a lower-dimensional Gaussian on the tangent plane, then projected downwards onto the manifold.

The lower-dimensional Gaussian density is used for the sampling distribution in a Metropolis ratio [4].

Tangent vectors are calculated from a SVD of P_θ

Two Theorems

The two approaches above are *equivalent*

Equivalent Distributions

The **Constrained GFD** and the **GFD with Parameterization** lead to the **same distribution** locally, and is invariant to the choice of **g.**

We can piece this local result to the whole space using a **partition of unity.**

Local Asymptotic Normality

In terms of the local parameter the Constrained GFD will approach a **Gaussian density on the manifold**, whenever it does so in the ambient space.

$$\int_{\mathcal{M}}\left|r_{\mathcal{M}}(\theta|\mathbf{y}) - \frac{\phi_{\hat\theta_n,nI(\theta_0)}(\theta)\left(\det\left|\nabla_t^2||g\circ\eta||^2\right|\right)^{-1/2}}{\int_{\mathcal{M}}\phi_{0,nI(\theta_0)}(s)\left(\det\left|\nabla_t^2||g\circ\eta||^2\right|\right)^{-1/2}d\lambda(\theta)}\right|d\lambda(\theta) \xrightarrow{P_{\theta_0}} 0$$

Local asymptotic normality is "inherited" when a GFD is constrained to a manifold!

Two Further Examples

Linear logsplines and the AR(1) Model

Density Estimation

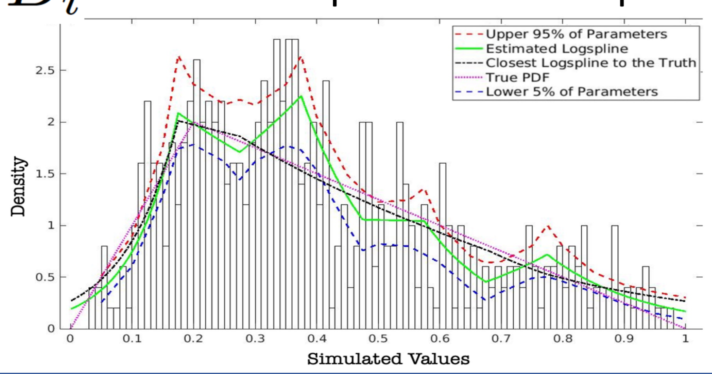
Consider the linear logsplines model

$$f(y|\boldsymbol{\theta}) = \exp\{\theta_1 B_1(y) + \cdots + \theta_d B_d(y)\}$$

Subject to the constraint that it is a probability density:

$$g(\boldsymbol{\theta}) = \log\left(\int_0^1 \exp\{\theta_1 B_1(y) + \cdots + \theta_d B_d(y)\}\,dy\right) = 0,$$

where B_i are linear B-splines. We sample from θ:



AR(1) Model

The MVN representation of an AR(1) model has cov matrix:

$$\Sigma = \frac{\sigma^2}{1-\rho^2}\begin{pmatrix} 1 & \rho & \rho^2 & \dots\rho^{n-1} \\ \rho & 1 & \rho & \dots\rho^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{pmatrix}$$

which is a general cov matrix subject to the constraints

$$g(\Sigma) = \left(\mathfrak{S}^1_{1,1}\dots\mathfrak{S}^1_{n-1,n-1}\,\mathfrak{S}^2_1\dots\mathfrak{S}^2_{n-2}\right)' = \mathbf{0},$$

where

$$\mathfrak{S}^1_{i,j} = \Sigma_{i,j} - \Sigma_{i+1,j+1}$$

$$\mathfrak{S}^2_k = \frac{\Sigma_{1,i+1}}{\Sigma_{1,i}} - \frac{\Sigma_{1,i+2}}{\Sigma_{1,i+1}}$$

$$1\leqslant i\leqslant j\leqslant n-1.$$

$$1\leqslant k\leqslant n-2.$$

Contact
Alexander C. Murph
University of NC at Chapel Hill
acmurph@live.unc.edu

Paper
[projected to be on arXiv by BFF 2022]

References

[1] Hannig, J., Iyer, H., Lai, R. C. S. and Lee, T. C. M. (2016) Generalized fiducial inference: A review and new results. Journal of the American Statistical Association, 111, 1346–1361.
[2] H. C. Andersen. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. Journal of Computational Physics, 52 (1):24 – 34, 1983. ISSN 0021-9991.
[3] Brubaker, M. A., Salzmann, M. and Urtasun, R. (2012) A Family of MCMC Methods on Implicitly Defined Manifolds. In Proc. of the 15th Intern. Conf. on AI and Statistics, vol. 22, 161–172.
[4] Zappa, E., Holmes-Cerfon, M. and Goodman, J. (2018) Monte carlo on manifolds: Sampling densities and integrating functions. Communications on Pure and Applied Mathematics, 71, 2609–2647.

Acknowledgements
The project was partially funded by NSF Award #1916115. Thank you also for the travel funding!