Statistica Sinica: Preprint
doi:10.5705/ss.202016.0017



# Bayesian nonparametric inference on the Stiefel manifold

Lizhen Lin, Vinayak Rao and David Dunson

*Department of Statistics and Data Sciences, The University of Texas at Austin, USA.*
*Department of Statistics, Purdue University, USA.*
*Department of Statistical Science, Duke University, USA.*

*Abstract:* The Stiefel manifold $V_{p,d}$ is the space of all $d \times p$ orthonormal matrices, with the $d-1$ hypersphere and the space of all orthogonal matrices constituting special cases. In modeling data lying on the Stiefel manifold, parametric distributions such as the matrix Langevin distribution are often used; however, model misspecification is a concern and it is desirable to have nonparametric alternatives. Current nonparametric methods are mainly Fréchet-mean based. We take a fully generative nonparametric approach, which relies on mixing parametric kernels such as the matrix Langevin. The proposed kernel mixtures can approximate a large class of distributions on the Stiefel manifold, and we develop theory showing posterior consistency. While there exists work developing general posterior consistency results, extending these results to this particular manifold requires substantial new theory. Posterior inference is illustrated on a dataset of near-Earth objects.

*Key words and phrases: Bayesian nonparametric, kernel mixture, matrix Langevin, orthonormal matrices, posterior consistency, Stiefel manifold, von Mises Fisher*

## 1. Introduction

Statistical analysis of matrices with orthonormal columns has diverse applications including principal components analysis, estimation of rotation matrices, as well as in analyzing orbit data of the orientation of comets and asteroids. Central to probabilistic models involving such matrices are probability distributions on the Stiefel manifold, the space of all $d \times p$ orthonormal matrices. Popular examples of parametric distributions are the matrix von Mises-Fisher distribution (Khatri and Mardia (1977); Hornik and Grün (2013)) (also known as the matrix Langevin (Chikuse (1993, 2003a, 2006))), and its generalization, the Bingham-von Mises-Fisher distribution (Hoff (2009)). Maximum likelihood estimation is often used in estimating the parameters, while recently Rao et al. (2016) proposed a sampling algorithm allowing Bayesian inference for such distributions.