Computational Statistics and Data Analysis 54 (2010) 1317–1327

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Finite-sample investigation of likelihood and Bayes inference for the symmetric von Mises–Fisher distribution

Melissa A. Bingham [a,*], Daniel J. Nordman [b], Stephen B. Vardeman [b]

[a] *University of Wisconsin-La Crosse, La Crosse, WI 54601, United States*
[b] *Iowa State University, United States*

ARTICLE INFO

Article history:
Received 15 March 2009
Received in revised form 17 October 2009
Accepted 23 November 2009
Available online 29 November 2009

ABSTRACT

We consider likelihood and Bayes analyses for the symmetric matrix von Mises–Fisher (matrix Fisher) distribution, which is a common model for three-dimensional orientations (represented by 3×3 orthogonal matrices with a positive determinant). One important characteristic of this model is a 3×3 rotation matrix representing the modal rotation, and an important challenge is to establish accurate confidence regions for it with an *interpretable* geometry for practical implementation. While we provide some extensions of one-sample likelihood theory (e.g., Euler angle parametrizations of modal rotation), our main contribution is the development of MCMC-based Bayes inference through non-informative priors. In one-sample problems, the Bayes methods allow the construction of inference regions with transparent geometry and accurate frequentist coverages in a way that standard likelihood inference cannot. Simulation is used to evaluate the performance of Bayes and likelihood inference regions. Furthermore, we illustrate how the Bayes framework extends inference from one-sample problems to more complicated one-way random effects models based on the symmetric matrix Fisher model in a computationally straightforward manner. The inference methods are then applied to a human kinematics example for illustration.

1. Introduction

Data points as three-dimensional rotation matrices arise in investigations in materials science (e.g., crystal orientations in metals Mackenzie, 1957; Bingham et al., 2009a) as well as studies of human kinematics (Rancourt et al., 2000), and the matrix Fisher is the most widely referenced distribution for modeling such observations (Downs, 1972). A large literature exists on likelihood methodology for this distribution, but there has been considerably less work on Bayes analyses. For example, advances in maximum likelihood inference for the matrix Fisher distribution have been made by Khatri and Mardia (1977) and Jupp and Mardia (1979), with further work done by Prentice (1986), Mardia and Jupp (2000), and Chikuse (2003). Chang and Rivest (2001) also developed *M*-estimation connected to likelihood inference. In terms of Bayes inference for the matrix Fisher distribution, Chang and Bingham (1996) outlined an approach for stipulating informative priors, while in an unpublished work, Camano-Garcia (unpublished) considered Gibbs samplers for general Langevin (or matrix Fisher) distributions.

Despite these developments and the clear popularity of the matrix Fisher distribution, relatively little appears to be known about the potential benefits of Bayes inference for this distribution or about the relative finite-sample performances of likelihood and Bayes methods for this model. For example, the Bayes development in Chang and Bingham (1996) targeted large-sample approximations of posterior distributions, often with informative priors amenable to such approximations. We

* Corresponding author.
E-mail address: bingham.meli@uwlax.edu (M.A. Bingham).

0167-9473/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2009.11.020

aim instead to explore a practical implementation of Bayes methods based on non-informative priors, under the common parametrization of the symmetric matrix Fisher used by Chang and Bingham (1996).

This matrix Fisher distribution is useful for modeling symmetric random errors around a fixed "central location" **S**, which is a 3×3 rotation matrix that constitutes an unusual model parameter but a key model summary in the existing literature. We show that Bayes methods offer computationally straightforward inference for this and, in even small samples, can produce inference regions with frequentist coverage rates that match nominal credible levels. Additionally, while a likelihood approach can provide a confidence region for **S** through a usual chi-square calibration, the likelihood function itself does not "order" the matrix-valued parameter space for **S** in a geometrically clear way that easily conveys the notion of inference set "size". As a consequence, likelihood confidence regions have mathematically convenient set-theoretic definitions, but do not clearly convey information about statistical precision. We illustrate, however, that non-informative Bayes regions for **S** may be constructed to have a simple geometrical structure that indicates precision (size) while retaining accurate frequentist coverages. In these ways, the Bayes methods can provide improved frequentist inference for the symmetric matrix Fisher distribution.

Beyond the one-sample problem, we also build and illustrate Bayes inference in one-way random effects models based on the symmetric matrix Fisher distribution. Our intent in this is to stimulate greater interest in Bayes methods for the Fisher model by demonstrating how such methods extend naturally from the one-sample situation to more complex inference scenarios. The Bayes framework may be more tractable than a purely likelihood approach for some problems and may open new inference possibilities for the matrix Fisher distribution.

Section 2 provides a preliminary framework, including a constructive description of the symmetric matrix Fisher distribution and a likelihood formulation. Although we focus on the matrix Fisher distribution in particular, the Bayes and likelihood methodologies presented here can apply to a larger class of distributions for random rotations which, as explained in Section 2, have the same kind of "constructive" definition and parametrization as the symmetric matrix Fisher. Asymptotic distributional results for one-sample likelihood inference are given based on an Euler angle representation of **S**, and simulations illustrate these approximations applied to the finite-sample distributions of likelihood statistics. Section 3 describes the Bayes methods for the one-sample problem using non-informative priors and a Metropolis–Hastings within Gibbs algorithm. Numerical studies compare the finite-sample properties of likelihood and Bayes inference regions for **S** where we impose a shape constraint on these regions to facilitate interpretation. Simulations indicate that the Bayes regions tend to have better coverage properties, even in samples as small as $n = 10$. Section 4 then examines the performance of Bayes methods for one-way random effects models built on the symmetric matrix Fisher distribution, while Section 5 illustrates the matrix methods with human kinematics data. Section 6 provides some conclusions.

2. The symmetric matrix Fisher distribution and one-sample likelihood

2.1. Constructive definition and model density

We first give a simple construction for generating random rotations from the symmetric Fisher model, from which the matrix density follows. Let Ω represent the set of 3×3 rotation matrices (orthogonal matrices that preserve the right hand rule). The symmetric matrix Fisher distribution may be characterized through a location parameter $\mathbf{S} \in \Omega$ and a spread parameter $\kappa > 0$, where the model itself (denoted by $\mathrm{F}(\mathbf{S}, \kappa)$ here) describes the deviation of random orientations $\mathbf{O} \in \Omega$ from a common "central" orientation $\mathbf{S} \in \Omega$. $\mathbf{O} \in \Omega$ from a $\mathrm{F}(\mathbf{S}, \kappa)$ model may be constructed as $\mathbf{O} = \mathbf{S} \cdot \mathbf{P}$, where a random perturbation

$$\mathbf{P} = \mathbf{U}\mathbf{U}^T + (\mathbf{I}_{3\times 3} - \mathbf{U}\mathbf{U}^T)\cos r + \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix} \sin r \in \Omega \tag{1}$$

is a rotation matrix built from two independent components: a unit vector $\mathbf{U}^T = (u_1, u_2, u_3) \in \mathbb{R}^3$ identified by a point uniformly distributed on the unit sphere and an angle $r \in (-\pi, \pi]$ distributed according to a density (with respect to the Lebesgue measure)

$$C(r|\kappa) = \frac{(1 - \cos r)\exp(2\kappa \cos r)}{2\pi (I_0(2\kappa) - I_1(2\kappa))}, \quad r \in (-\pi, \pi]; \tag{2}$$

above I_i denotes the modified Bessel function of order i and the parameter $\kappa \in (0, \infty)$ controls the spread of the angle density (2), which is symmetric around zero. The matrix **P** represents the positions of coordinate axes in $\mathbb{R}^3$ (denoted by the 3×3 identity matrix $\mathbf{I}_{3\times 3}$) after spinning the $\mathbb{R}^3$-frame around the random axis $\mathbf{U} \in \mathbb{R}^3$ by the random angle r. Note that a small $|r|$ value in (1) entails a small rotational deviation **P** from $\mathbf{I}_{3\times 3}$ (e.g., $r = 0$ implies $\mathbf{P} = \mathbf{I}_{3\times 3}$) and, since κ controls the spread or concentration of r around 0, this parameter also controls the variation of a $\mathrm{F}(\mathbf{S}, \kappa)$ observation $\mathbf{O} = \mathbf{SP}$ from the location parameter $\mathbf{S} \in \Omega$.

To understand how the spread component κ in the density (2) of r translates into "spread" for the $\mathrm{F}(\mathbf{S}, \kappa)$ distribution, consider Table 1. Here, $\Delta_1(\kappa)$ is the distributional median of $|r|$, where $|r|$ is sometimes referred to as a "misorientation angle" with density $2 \cdot C(|r||\kappa)$, and $\Delta_2(\kappa)$ represents the distributional median of the maximum angle between standard coordinate axes $\mathbf{S} = \mathbf{I}_{3\times 3}$ and the same axes rotated by a $\mathrm{F}(\mathbf{I}_{3\times 3}, \kappa)$ observation (i.e., the maximum angle over 3 rotated