Simulation of random rotation matrices

John T. Kent and Asaad M. Ganeiber

Abstract Directional distributions are playing an increasing role as building blocks in sophisticated geometric statistical models. Inference from such models is often carried out by MCMC. Hence it is important to have efficient methods of simulation for the underlying directional distributions. In this paper we survey some existing methods and describe a method for the matrix Fisher distribution for 3×3 rotation matrices which is based on a new acceptance-rejection simulation method for the Bingham distribution.

Key words: matrix Fisher distribution, acceptance-rejection simulation, angular central Gaussian distribution, Bingham distribution

1 Introduction

Directional data analysis is concerned with statistical analysis on various non-Euclidean manifolds, starting with circle and the sphere, and extending to related manifolds (Mardia and Jupp, 2000). Directional distributions can used as building blocks in more sophisticated statistical models which are studied using MCMC methods. For example, Green and Mardia (2006) used the matrix Fisher distribution in a Bayesian model to align two configurations of points in $\mathbb{R}^3$ in an unlabelled version of shape analysis, and they applied the model to a problem of protein alignment in bioinformatics. Hence there is a need to develop simulation methods for directional distributions which are efficient over a wide range of concentration parameters. In this paper we focus on the simulation of the matrix Fisher distribution

John T. Kent
University of Leeds, Leeds LS2 9JT, UK. e-mail: j.t.kent@leeds.ac.uk

Asaad M. Ganeiber
University of Leeds, Leeds LS2 9JT, UK. e-mail: asaad@maths.leeds.ac.uk

Table 1 Some common distributions on directional spaces

Space	Notation	Distributions
circle	S_1	von Mises, wrapped Cauchy
sphere	S_p	Fisher ($p = 2$) and von-Mises-Fisher ($p \geq 1$), Fisher-Bingham
real projective space	$\mathbb{R}P_p$	Bingham, angular central Gaussian
special orthogonal group	$SO(q)$	matrix Fisher

on the space of 3×3 rotations using a new acceptance-rejection method to simulate the Bingham distribution.

2 Directional distributions

Table 1 gives some of the common spaces associated with directional data analysis, together with the main distributions.

The sphere $S_p = \{x \in \mathbb{R}^{p+1} : \; x^T x = 1\}$ represents the space of "directions" in $\mathbb{R}^{p+1}$. Real projective space consists of the "axes" or "unsigned directions" $\pm x$. In some sense this space is half of a sphere; it can also be represented as the space of rank 1 projection matrices,

$$\mathbb{R}P_p = \{P \in \mathbb{R}^{(p+1)\times(p+1)} : \; P = P^T, \; P^2 = P, \; \mathrm{tr}(P) = 1\}. \tag{1}$$

A rank one projection matrix can be written as $P = xx^T$ where x is a unit vector. The special orthogonal group of $q \times q$ rotation matrices is defined by

$$SO(q) = \left\{R \in \mathbb{R}^{q\times q} : \; \det R = 1, \; R^T R = I_q\right\},$$

On each of these spaces there is a unique uniform distribution which is invariant under rotations. Further each of these spaces is naturally embedded in a Euclidean space. A natural "linear-exponential" family of distributions can be generated by letting the density (with respect to the uniform measure) be proportional to the exponential of a linear function of the Euclidean variables. This construction generates the first named distribution in each of the four rows of the table above. It should be noted that the Bingham distribution, whose log density is linear in $P = xx^T$ in (1), can also be viewed as a distribution on the sphere whose log density is quadratic in x.

3 The matrix Fisher distribution

The linear-exponential family on $SO(p)$ is known as the matrix Fisher distribution, with density

$$f(X) = c_F \exp\left\{\mathrm{tr}\left(F^T X\right)\right\}, \quad X \in SO(p),$$

with respect to the underlying invariant Haar measure. This density was introduced by Khatri and Mardia (1977); it is unimodal about a fixed rotation matrix determined by the $p \times p$ parameter matrix F.

Now specialize to the case $q = 3$. A matrix in $X \in SO(3)$ can be written in the form $X = H_{23}(\phi)H_{13}(\theta)H_{12}(\psi)$, where for $1 \le i < j \le 3$, $H_{ij}(\theta)$ denotes a 3×3 matrix which looks like an identity matrix except for values $\cos\theta$ in locations (i,i) and (j,j), and values $\sin\theta$ and $-\sin\theta$ in locations (i,j) and (j,i). Thus X is constructed as a product of three two-dimensional rotations about each of the coordinate axes in turn. The angles ϕ, θ, ψ are known as Euler angles. They lie in ranges, $0 \le \phi, \psi < 2\pi$ and $-\pi/2 \le \theta \le \pi/2$. In these coordinates the underlying Haar measure can be represented as

$$[\mathrm{d}X] = \cos\theta \mathrm{d}\theta \mathrm{d}\phi \mathrm{d}\psi.$$

Note the presence of the $\cos\theta$ factor, which arises because small circles of constant latitude have a smaller circumference near the poles than near the equator.

The matrix Fisher distribution reduces to the uniform distribution if $F = 0$ and becomes more concentrated about its modal value as the overall concentration $||F|| = \{\mathrm{tr}(F^T F)\}^{1/2}$ increases. For theoretical purposes it suffices to limit attention to the diagonal case $F = \Delta = \mathrm{diag}(\delta_j)$, where $\delta_1 \ge \delta_2 \ge |\delta_3|$. As the concentration increases, the distribution becomes concentrated near $\theta = \phi = \psi = 0$, and asymptotically, $f(X)$ becomes a trivariate normal distribution,

$$f(X) \propto \exp\left\{-\frac{1}{2}\left[(\delta_1 + \delta_3)\,\theta^2 + (\delta_1 + \delta_2)\,\phi^2 + (\delta_2 + \delta_3)\,\psi^2\right]\right\},$$

with respect to Lebesgue measure $\mathrm{d}\theta\mathrm{d}\phi\mathrm{d}\psi$.

4 Simulation

When developing acceptance-rejection simulation methods for directional distributions, there are several issues to consider:

- the need for good efficiency for a wide range of concentration parameters, from uniform to highly concentrated. In similar problems on $\mathbb{R}^p$, the task is simpler when distributions are closed under affine transformations; in such cases it is often sufficient to consider just a single standardized form of the distribution.
- the need for a tractable envelope distribution.
- the presence of trigonometric factors in the base measure.

Efficient acceptance-rejection methods are available for the simpler directional distributions, most notably the Best-Fisher method (Best and Fisher, 1979) for the von Mises distribution. For the more complicated distributions, several MCMC algorithms have recently been proposed, e.g. Green and Mardia (2006); Kume and

Walker (2006); Hoff (2009). However, acceptance-rejection methods with reasonable acceptance probabilities are to be preferred when available. The following steps outline the way a new acceptance-rejection simmulation algorithm for the Bingham distribution can be applied to the matrix Fisher distribution.

- A classic result from differential geometry states that the space SO(3) can be identified with real projective space $\mathbb{R}P_3$ under a one-to-one mapping, or equivalently with the unit sphere S_3 in $\mathbb{R}^4$ under a one-to-two mapping. Each rotation matrix on SO(3) maps to two antipodal points on this unit sphere. This identification is limited to the case $q = 3$. There does not seem to be any useful analogue for $SO(q),\ q > 3$.
- The matrix Fisher distribution on $SO(3)$ corresponds to the Bingham distribution on S_3.
- The PhD thesis of the second author gives a new method to simulate from the Bingham distribution on S_p for any $p \geq 1$ using an acceptance-rejection algorithm with the angular central Gaussian distribution as an envelope.
- The angular central Gaussian distribution on S_p is very simple to simulate. Given a $(p+1) \times (p+1)$ covariance matrix Σ, simulate $y \sim N_{p+1}(0, \Sigma)$ and set $z = y/||y||$. Given the parameters of a Bingham distribution, it is possible to determine a choice of Σ to give a good envelope.
- The use of an angular central Gaussian envelope for the Bingham distribution is closely related to the use of a multivariate Cauchy density as an envelope for simulating a multivariate normal distribution.
- the acceptance ratio is typically at least 45% for a wide range of parameters. This value is very reasonable for practical purposes.

References

Best DJ, Fisher NI (1979) Efficient simulation of the von Mises distribution. J Appl Stat 28:152–157

Green PJ, Mardia KV (2006) Bayesian alignment using hierarchical models, with applications in protein bioinformatics. Biometrika 93:235–254

Hoff PD (2009) Simulation of the matrix Bingham-von Mises-Fisher distribution with application to multivariate and relational data. J Comput Graph Stat 18:438–456

Khatri CG, Mardia KV (1977) The von Mises-Fisher matrix distribution in orientation statistics. J Roy Stat Soc B 39:95–106

Kume A, Walker SG (2006) Sampling from compositional and directional distributions. Stat Comput 16:261–265

Mardia KV, Jupp PE (2000) Directional Statistics. Wiley