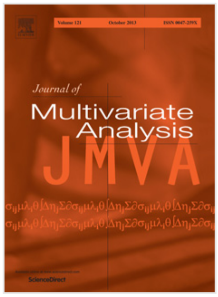
Accepted Manuscript



Copulae on products of compact Riemannian manifolds

P.E. Jupp

PII: S0047-259X(15)00112-8
DOI: http://dx.doi.org/10.1016/j.jmva.2015.04.008
Reference: YJMVA 3923

To appear in: *Journal of Multivariate Analysis*

Received date: 7 October 2014

Copulae on products of compact Riemannian manifolds

P. E. Jupp

School of Mathematics and Statistics, University of St Andrews, St Andrews KY16 9SS, UK

Abstract

One standard way of considering a probability distribution on the unit n-cube, $[0,1]^n$, due to Sklar (1959) [A. Sklar, Fonctions de répartition à n dimensions et leur marges, Publ. Inst. Statist. Univ. Paris 8 (1959) 229-231], is to decompose it into its marginal distributions and a copula, i.e. a probability distribution on $[0,1]^n$ with uniform marginals. The definition of copula was extended by Jones *et al.* (2014) [M.C. Jones, A. Pewsey, S. Kato, On a class of circulas: copulas for circular distributions, Ann. Inst. Statist. Math., to appear] to probability distributions on products of circles. This paper defines a copula as a probability distribution on a product of compact Riemannian manifolds that has uniform marginals. Basic properties of such copulae are established. Two fairly general constructions of copulae on products of compact homogeneous manifolds are given; one is based on convolution in the isometry group, the other using equivariant functions from compact Riemannian manifolds to their spaces of square integrable functions. Examples illustrate the use of copulae to analyse bivariate spherical data and bivariate rotational data.

Keywords: Bivariate, Convolution, Homogeneous manifold, Markov process, Uniform distribution, Uniform scores
2013 MSC: 62H11

1. Introduction

An n-copula is a probability distribution on the unit n-cube, $[0,1]^n$, which has uniform marginals. An n-copula and n probability distributions on $\mathbb{R}$

Email address: pej@st-andrews.ac.uk (P. E. Jupp)

together determine a probability distribution on $\mathbb{R}^n$ with n-dimensional distribution function, H, defined by

$$H(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)), \tag{1}$$

where C is the n-dimensional distribution function of the copula and $F_1, \ldots, F_n$ are the cumulative distribution functions of the distributions on $\mathbb{R}$. Sklar's theorem [27] states that (i) any n-dimensional distribution function, H, has the form (1) for some C and $F_1, \ldots, F_n$, (ii) if $F_1, \ldots, F_n$ are continuous then C is unique. Thus copulae can be used (a) to generate families of multivariate distributions by coupling together distributions with given marginals, (b) to decompose any multivariate distribution into (i) its marginal distributions and (ii) a copula, which describes the dependence structure. Comprehensive accounts of copulae are given in the books of Joe [9] and Nelsen [23].

The use of copulae in (1) to relate multivariate probability distributions to their marginal distributions has been generalised from $[0, 1]^n$ to products of Polish product spaces by Scarsini [26]. This approach requires the choice of linearly ordered increasing classes of subsets of each of the spaces in the product. No such classes are required in the approach taken here.

Copulae on the circle, called 'circulae', have been developed by Jones, Pewsey & Kato [12], who pay particular attention to a method of constructing circulae using convolution, special cases of which had appeared in [10], [8], [11] and [28]. The aim of this paper is to extend the idea of circula to a concept of copula on compact Riemannian manifolds, which are the main sample spaces considered in directional statistics.

Section 2 concerns uniform distributions on compact Riemannian manifolds and considers the existence and uniqueness of homeomorphisms generalising probability integral transformations, in that they transform given probability distributions into the uniform distribution. Copulae on compact Riemannian manifolds are defined in Section 3, and various equivalent formulations are given. Section 4 considers two general constructions of copulae on compact Lie groups and their coset spaces. One of these extends the convolution construction in [12] of circulae; the other is based on the Sobolev mappings used in directional statistics. Examples involving paired data on the sphere and on the rotation group, $SO(3)$, are given in Section 5.

2. Uniform distributions and uniform scores

2.1. Uniform distributions

Let $\mathcal{X}$ be a compact Riemannian manifold. Then the Riemannian metric determines the volumes of infinitesimal cubes, and so equips $\mathcal{X}$ with a unique uniform probability measure, $\nu_{\mathcal{X}}$. If $\phi : \mathcal{X} \to \mathcal{X}$ is continuous and μ is a probability distribution on $\mathcal{X}$ then $\phi_*\mu$ denotes the image measure given by $(\phi_*\mu)(A) = \mu(\phi^{-1}(A))$ for all measurable subsets A of $\mathcal{X}$.

Proposition 1. *For any probability distribution μ on $\mathcal{X}$ such that the density of μ with respect to $\nu_{\mathcal{X}}$ is continuous and positive, there is a homeomorphism ϕ of $\mathcal{X}$ such that $\phi_*\mu = \nu_{\mathcal{X}}$. If $\mathcal{X} = [0,1]$ then there is a unique such ϕ with $\phi(0) = 0$. If $\mathcal{X}$ is the circle, S^1, and a base-point θ_0 and an orientation of S^1 are specified then there is a unique orientation-preserving such ϕ with $\phi(\theta_0) = \theta_0$.*

Proof. The homeomorphism ϕ can be constructed using the probability integral transformation in coordinate neighourhoods, as in the first of the proofs in [22] for the case in which μ has a smooth density. □

For $\mathcal{X} = S^1$, Jones *et al.* [12] point out that it is convenient to take θ_0 to be the mode (assumed unique) of μ. If the dimension of $\mathcal{X}$ is greater than 1 then the homeomorphism ϕ is not unique, even if ϕ is required to keep some given point fixed. The following proposition shows that, in some cases, there is a canonical choice of ϕ. First recall that, if $\mathcal{X}$ is compact, connected and without boundary, and x_0 is any point in $\mathcal{X}$ then the exponential map from the tangent space, $T\mathcal{X}_{x_0}$, at x_0 into $\mathcal{X}$ defines a maximal system of Riemannian normal coordinates around x_0 as follows. The inverse of this coordinate system maps the open set $\{(t, \mathbf{v}) : 0 \leq t < r_{\mathbf{v}}, \mathbf{v} \in T_1\mathcal{X}_{x_0}\}$ diffeomorphically onto an open set U of $\mathcal{X}$ by $(t, \mathbf{v}) \mapsto \exp(t\mathbf{v})$. Here $T_1\mathcal{X}_{x_0}$ denotes the set of unit tangent vectors at x_0 and

$$r_{\mathbf{v}} = \sup\{t : \text{there is a unique minimising geodesic from } x_0 \text{ to } \exp(t\mathbf{v})\}.$$

See, e.g. Sections VII.6 and VII.7 of [2] or Section II.C of [6]. It follows from nullity of the cut locus (see, e.g. Lemma 3.96 of [6]) that $\mathcal{X} \setminus U$ has measure zero. Thus absolutely continuous probability distributions on $\mathcal{X}$ can be identified with absolutely continuous probability distributions on $\{(t, \mathbf{v}) : 0 \leq t < r_{\mathbf{v}}, \mathbf{v} \in T_1\mathcal{X}_{x_0}\}$. In particular, such a distribution induces a marginal distribution on $T_1\mathcal{X}_{x_0}$.

Proposition 2. *Let μ be a probability distribution on $\mathcal{X}$ such that the density of μ with respect to $\nu_{\mathcal{X}}$ is continuous and positive. Let x_0 be a point of $\mathcal{X}$. If the distribution induced by μ on the unit tangent sphere $T_1\mathcal{X}_{x_0}$ is the same as that induced by ν_X then there are (i) an open set U in $\mathcal{X}$ such that $\mathcal{X} \setminus U$ has measure zero, (ii) a unique homeomorphism $\phi : U \to U$, such that $\phi_*\mu = \nu_{\mathcal{X}}$, and ϕ sends each geodesic ray in U that starts at x_0 into itself.*

Proof. The homeomorphism ϕ is constructed using the probability integral transform along each geodesic. The open set U is given by $U = \{(t, \mathbf{v}) : 0 \le t < r_{\mathbf{v}}, \mathbf{v} \in T_1\mathcal{X}_{x_0}\}$, and ϕ is defined by $\phi(\exp(t\mathbf{v})) = \exp(u\mathbf{v})$, where $F_{\mathbf{v}}(t) = F_{\mathbf{v},0}(u)$, $F_{\mathbf{v}}$ and $F_{\mathbf{v},0}$ being the cumulative distribution functions of t conditional on $\mathbf{v}$ given by μ and $\nu_{\mathcal{X}}$, respectively. □

Example 1. *Let $\mathcal{X}$ be S^2 and μ be the Fisher distribution with mean direction $\boldsymbol{\mu}$ and concentration κ. Then calculation shows that the homeomorphism ϕ of Proposition 2 is $\phi(\mathbf{x}) = u\boldsymbol{\mu} + \sqrt{(1-u^2)/(1-t^2)}\,(\mathbf{I}_3 - \boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}})\,\mathbf{x}$, where $t = \mathbf{x}^{\mathrm{T}}\boldsymbol{\mu}$, $u = (2e^{\kappa t} - e^{\kappa} - e^{-\kappa})\,/\,(e^{\kappa} - e^{-\kappa})$ and $\mathbf{I}_3$ denotes the 3×3 identity matrix.*

Example 2. *Let $\mathcal{X}$ be $SO(3)$ and μ be the matrix Fisher distribution with density proportional to $\exp\{\mathrm{tr}\,(\kappa\mathbf{X}^{\mathrm{T}}\mathbf{M})\}$ for $\kappa \ge 0$ and $\mathbf{X}, \mathbf{M}$ in $SO(3)$. Since the geodesics through $\mathbf{I}_3$ in $SO(3)$ are rotations of constant speed about fixed axes, calculation shows that for the homeomorphism ϕ of Proposition 2, $\mathbf{M}^{\mathrm{T}}\mathbf{X}$ and $\mathbf{M}^{\mathrm{T}}\phi(\mathbf{X})$ have the same rotation axis, and that the rotation angle, u, of $\mathbf{M}^{\mathrm{T}}\phi(\mathbf{X})$is related to the rotation angle, t, of $\mathbf{M}^{\mathrm{T}}\mathbf{X}$ by*

$$\tilde{F}_0(u)/\tilde{F}_0(\pi) = \tilde{F}_\kappa(t)/\tilde{F}_\kappa(\pi),$$

where $\tilde{F}_\kappa(\theta) = \int_0^\theta e^{4\kappa\cos^2(\omega/2)}\sin^2(\omega/2)d\omega$.

2.2. Uniform scores

The following discrete version of Proposition 2 can be useful in data analysis. Let x_0 be a base-point of $\mathcal{X}$ and suppose that the uniform distribution is symmetric about x_0, in the sense that, in the normal coordinates $(t, \mathbf{v})$, $\mathbf{v}$ is uniformly distributed on $T_1\mathcal{X}_{x_0}$ and t is independent of $\mathbf{v}$. (Riemannian manifolds with this property include spheres, projective spaces and the rotation group $SO(3)$ with their usual metrics.) For almost all points $x_1, \ldots, x_n$, each x_i can be written as $x_i = \exp(t_i\mathbf{v}_i)$ for some $(t_i, \mathbf{v}_i)$ with $\mathbf{v}_i \in T_1\mathcal{X}_{x_0}$ and $0 \le t_i < r_{\mathbf{v}_i}$. For $i = 1, \ldots, n$, define $j(i)$ by $t_i = t_{(j(i))}$, the $j(i)$th smallest order statistic of $t_1, \ldots, t_n$, and define u_i by $F_0(u_i) = \{j(i) - 1/2\}/n$, where F_0 denotes the cumulative distribution function of t conditional on $\mathbf{v}$ under the uniform distribution, $\nu_{\mathcal{X}}$. The *uniform scores*, $\tilde{x}_1, \ldots, \tilde{x}_n$ are defined by $\tilde{x}_i = \exp(u_i \mathbf{v}_i)$. If $x_1, \ldots, x_n$ is a random sample from a distribution on $\mathcal{X}$ which is symmetric about x_0 then as $n \to \infty$ the empirical distribution given by $\{\tilde{x}_1, \ldots, \tilde{x}_n\}$ tends to the uniform distribution on $\mathcal{X}$. In the case $\mathcal{X} = S^1$, $\tilde{x}_1, \ldots, \tilde{x}_n$ are very similar to the usual uniform scores that are used in some two-sample tests on the circle, as in [21, Section 8.3.1].

Given a base-point (x_0, y_0) in $\mathcal{X} \times \mathcal{Y}$ where the uniform distributions on $\mathcal{X}$ and $\mathcal{Y}$ are symmetric about x_0 and y_0, respectively, a simple extension of the above construction transforms sets of points $(x_1, y_1), \ldots, (x_n, y_n)$ into uniform scores, $(\tilde{x}_1, \tilde{y}_1), \ldots, (\tilde{x}_n, \tilde{y}_n)$. If $(x_1, y_1), \ldots, (x_n, y_n)$ is a random sample from a distribution on $\mathcal{X} \times \mathcal{Y}$ having marginals that are symmetric about x_0 and y_0, respectively, then as $n \to \infty$ the empirical distribution given by $\{(\tilde{x}_1, \tilde{y}_1), \ldots, (\tilde{x}_n, \tilde{y}_n)\}$ tends to a distribution on $\mathcal{X} \times \mathcal{Y}$ having unform marginals, i.e. a copula in the sense of Section 3 below. In the case $\mathcal{X} = \mathcal{Y} = S^1$, $\tilde{x}_1, \ldots, \tilde{x}_n$ and $\tilde{y}_1, \ldots, \tilde{y}_n$ are similar to the usual uniform scores that are used in the uniform scores Sobolev tests of independence considered in [18, Section 3].

3. Copulae on compact Riemannian manifolds

Let $\mathcal{X}$ and $\mathcal{Y}$ be compact Riemannian manifolds. A *copula* (more precisely, a 2-copula) on $\mathcal{X} \times \mathcal{Y}$ is a probability distribution on $\mathcal{X} \times \mathcal{Y}$ that has uniform marginal distributions on $\mathcal{X}$ and $\mathcal{Y}$. More generally, an *n-copula* on a product $\prod_{i=1}^n \mathcal{X}_i$ of compact Riemannian manifolds is a probability distribution on $\prod_{i=1}^n \mathcal{X}_i$ with uniform marginal distribution on each $\mathcal{X}_i$. For simplicity, only 2-copulae will be considered here.

The following proposition follows immediately from Proposition 1.

Proposition 3. *Let μ be a probability distribution on $\mathcal{X} \times \mathcal{Y}$ having marginal densities with respect to $\nu_{\mathcal{X}}$ and $\nu_{\mathcal{Y}}$ that are continuous and positive. Then there are homeomorphisms ϕ_X of $\mathcal{X}$ and ϕ_Y of $\mathcal{Y}$ such that $(\phi_X \times \phi_Y)_* \mu$ is a copula.* □

In general, the copula in Proposition 3 is not unique.

For any compact Riemannian manifold $\mathcal{X}$, let $\mathcal{P}(\mathcal{X})$ denote the set of probability distributions on $\mathcal{X}$ that are mutually absolutely continuous with respect to $\nu_{\mathcal{X}}$. The following Proposition shows that copulae can be considered from several equivalent viewpoints. The proof is straightforward.

Proposition 4. *For a probability distribution μ on $\mathcal{X} \times \mathcal{Y}$ with marginals that are mutually absolutely continuous with respect to $\nu_{\mathcal{X}}$ and $\nu_{\mathcal{Y}}$, let $\mu_{\mathcal{X}}$ and $\mu_{\mathcal{Y}}$ denote the mappings (Markov kernels) from $\mathcal{X}$ to $\mathcal{P}(\mathcal{Y})$ and from $\mathcal{Y}$ to $\mathcal{P}(\mathcal{X})$ for which $\mu_{\mathcal{X}}(x)$ and $\mu_{\mathcal{Y}}(y)$ are the conditional probability distributions on $\mathcal{Y}$ given x and on $\mathcal{X}$ given y, respectively. Then the maps $\mu \mapsto \mu_{\mathcal{X}}$ and $\mu \mapsto \mu_{\mathcal{Y}}$ give one-to-one correspondences between*

(a) the set of copulae on $\mathcal{X} \times \mathcal{Y}$ that are in $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$,

(b) the set of Markov kernels from $\mathcal{X}$ to $\mathcal{Y}$ that send $\nu_{\mathcal{X}}$ to $\nu_{\mathcal{Y}}$, i.e.

$$\int_{\mathcal{X}} \mu_{\mathcal{X}}(x) d\nu_{\mathcal{X}}(x) = \nu_{\mathcal{Y}},$$

(c) the set of Markov kernels from $\mathcal{Y}$ to $\mathcal{X}$ that send $\nu_{\mathcal{Y}}$ to $\nu_{\mathcal{X}}$.

If $\mathcal{Y} = \mathcal{X}$ then the identification of a Markov process with its transition Markov kernel and stationary distribution gives a one-to-one correspondence between any of (a), (b) or (c) and

(d) the set of discrete-time Markov processes on $\mathcal{X}$ with uniform stationary distribution. □

Using the correspondence between (a) and (b), composition of uniformity-preserving Markov kernels from $\mathcal{X}$ to $\mathcal{X}$ can be translated into a concept of composition of copulae on $\mathcal{X} \times \mathcal{X}$. For $\mathcal{X} = \mathbb{R}$ this was introduced by Darsow et al. [3]; see Section 6.4 of [23].

4. Copulae on compact homogeneous manifolds

4.1. Mixed translation copulae on compact Lie groups

Let G be a compact Lie group. The inversion map $x \mapsto x^{-1}$ and, for any elements z_1, z_2 of G, the 2-sided translation $x \mapsto z_1^{-1} x z_2$ are isometries and so preserve the uniform distribution. It follows that mixtures of these transformations have the same property. Thus, if X, Z_1, Z_2 are random variables on G with X uniformly distributed and independent of (Z_1, Z_2) then the distributions of $(X, Z_1^{-1} X Z_2)$ and $(X, Z_1^{-1} X^{-1} Z_2)$ are copulae on $G \times G$, called *positive mixed-translation copulae* and *negative mixed-translation copulae*, respectively. The distribution of (Z_1, Z_2) is called the *binding distribution*. If Z is in the centre of G, i.e. commutes with all elements of G, then $(ZZ_1)^{-1}X(ZZ_2) = Z_1^{-1}XZ_2$ and $(ZZ_1)^{-1}X^{-1}(ZZ_2) = Z_1^{-1}X^{-1}Z_2$.

If Z_1 and Z_2 are independent with densities f_1 and f_2, respectively, then the distributions of $(X, Z_1^{-1}XZ_2)$ and $(X, Z_1^{-1}X^{-1}Z_2)$ are called *positive convolution copulae* and *negative convolution copulae*, and have densities

$$\int_G f_1(xzy^{-1})f_2(z)d\nu_G(z) = \int_G f_1(u)f_2(x^{-1}uy)d\nu_G(u) \tag{2}$$

and

$$\int_G f_1(x^{-1}zy^{-1})f_2(z)d\nu_G(z) = \int_G f_1(u)f_2(xuy)d\nu_G(u), \tag{3}$$

respectively. For $G = S^1$, this construction is that considered by Wehrly & Johnson [28] and Jones et al. [12]. Special cases of the corresponding stationary Markov processes were given in models 3.1 and 3.4 of [1]. Taking $Z_2 = Z_1$ gives a random conjugation of X.

Two useful classes of convolution copulae are the *left convolution copulae*, for which $Z_2 = e$, and the *right convolution copulae*, for which $Z_1 = e$, where e denotes the identity element of G. Tests for a distribution on $G \times G$ to be a left or right convolution copula can be obtained from tests of uniformity. If the distribution of (X, Y) is a left (respectively, right) positive convolution copula then $(X, X^{-1}Y)$ (respectively, (X, YX^{-1})) is uniformly distributed on $G \times G$. Similarly, if the distribution of (X, Y) is a left (or right) negative convolution copula then (X, XY) (respectively, (X, YX)) is uniformly distributed on $G \times G$.

4.2. Convolution copulae on coset spaces

For closed subgroups K_1 and K_2 of a compact Lie group G, the corresponding left, right, and double coset spaces are $G/K_2 = \{xK_2 : x \in G\}$, $K_1\backslash G = \{K_1x : x \in G\}$ and $K_1\backslash G/K_2 = \{K_1xK_2 : x \in G\}$, respectively. Probability distributions on G/K_2, $K_1\backslash G$ and $K_1\backslash G/K_2$ can be identified with probability distributions on G that are (right) K_2-invariant, (left) K_1-invariant, and (left-right) $(K_1 \times K_2)$-invariant, respectively. Using these identifications, convolution of distributions on coset spaces can be defined as convolution of the corresponding invariant distributions on G. In the cases (a) $G = SO(p)$, $K_1 = \{e\}$, $K_2 = SO(p-2)$, (b) $G = SO(p)$, $K_1 = SO(p-2)$, $K_2 = SO(p-2)$, explicit descriptions of the convolutions on the coset spaces (a) S^{p-1} and (b) $[-1, 1]$ are given in Sections 9 and 6 of [19].

The above definition of convolution on coset spaces enables the following extension of convolution copulae from Lie groups to coset spaces. Let K_1, K_2, K_3, K_4 be closed subgroups of G. Denote by π_{12} and π_{34} the maps $x \mapsto K_1 x K_2$ and $x \mapsto K_3 x K_4$ from G to $K_1\backslash G/K_2$ and $K_3\backslash G/K_4$, respectively. Let X, Z_1, Z_2 be random variables on G with X uniformly distributed, (Z_1, Z_2) being independent of X and having any distribution on $G \times G$ for which the distributions of Z_1 and Z_2 are $(K_1 \times K_2)$-invariant and $(K_3 \times K_4)$-invariant, respectively. Then simple invariance arguments show that the distributions of $(\pi_{12}(X), \pi_{34}(Z_1^{-1} X Z_2))$ and $(\pi_{12}(X), \pi_{34}(Z_1^{-1} X^{-1} Z_2))$ are copulae on $(K_1\backslash G/K_2) \times (K_3\backslash G/K_4)$, called *convolution copulae* An important case is that in which $K_1 = K_3 = \{e\}$, so that these copulae are on $G/K_2 \times G/K_4$. The compact manifolds of the form G/K are the compact homogeneous manifolds, i.e. those on which the group of isometries acts transistively. If the density f_1 in (2) is $(K_1 \times K_3)$-invariant and f_2 is $(K_2 \times K_4)$-invariant then (2) is a copula on $(K_1\backslash G/K_2) \times (K_3\backslash G/K_4)$. If f_1 is $(K_2 \times K_3)$-invariant and f_2 is $(K_1 \times K_4)$-invariant then (3) is a copula on $(K_1\backslash G/K_2) \times (K_3\backslash G/K_4)$. For S^1 and S^2, examples of this construction were given in models 2.1, 2.4, 3.1 and 3.4 of [1].

4.3. Sobolev copulae on homogeneous compact manifolds

For a compact Riemannian manfold $\mathcal{X}$, let E_k $(k = 1, 2, \ldots)$ be the space of eigenfunctions of the Laplacian corresponding to the kth non-zero eigenvalue. Then the E_k are orthogonal finite-dimensional subspaces of the space $L^2(\mathcal{X})$ of square-integrable functions on $\mathcal{X}$. There are canonical maps $\mathbf{t}_k$ $(k = 1, 2, \ldots)$ of $\mathcal{X}$ into E_k, given by

$$\mathbf{t}_k(x) = \sum_{i=1}^{\dim\, E_k} f_i(x) f_i,$$

where $\{f_i : 1 \leq i \leq \dim\, E_k\}$ is any orthonormal basis of E_k. If $a_1, a_2, \ldots$ is a sequence of real numbers such that

$$\sum_{k=1}^{\infty} a_k^2 \dim\, E_k < \infty$$

then

$$x \mapsto \mathbf{t}(x) = \sum_{k=1}^{\infty} a_k \mathbf{t}_k(x) \tag{4}$$

defines a mapping $\mathbf{t}$ of $\mathcal{X}$ into $L^2(\mathcal{X})$. Such mappings $\mathbf{t}$ are the basis of Sobolev tests of uniformity [7, 15], several-sample tests [29], tests of symmetry [17], tests of independence [18], and tests of goodness of fit [14]. They can also be used as follows to construct copulae.

Proposition 5. *Let $\mathcal{X} = G/K$ be a compact homogeneous manifold, $\mathbf{t} : \mathcal{X} \to L^2(\mathcal{X})$ be a mapping of the form (4), h be a continuous real-valued function on $\mathbb{R}$ and μ be a probability distribution on G. For any real κ, define the probability density $f(\cdot, \cdot; \kappa, \mathbf{t}, \mu)$ on $\mathcal{X} \times \mathcal{X}$ by*

$$f(x, y; \kappa, \mathbf{t}, \mu) = c(\kappa) \int_G h(\kappa \langle \mathbf{t}(x), \mathbf{t}(gy) \rangle) d\mu(g), \tag{5}$$

where $c(\kappa)$ is a normalising constant and $\langle \cdot, \cdot \rangle$ denotes the inner product on $L^2(\mathcal{X})$. Then the distribution with density (5) is a copula on $\mathcal{X} \times \mathcal{X}$, called the Sobolev copula *given by $\mathbf{t}, h, \kappa$ and μ.*

If $\mathcal{X}$ is a Lie group, G, then the densities (5) can can be generalised to copula densities of the form

$$c(\kappa) \int_{G \times G} h(\kappa \langle \mathbf{t}(x), \mathbf{t}(g_1 y g_2^{-1}) \rangle) d\omega(g_1, g_2), \tag{6}$$

where ω is a probability distribution on $G \times G$.

PROOF. For γ in G,

$$\langle \mathbf{t}(\gamma x), \mathbf{t}(gy) \rangle = \langle \mathbf{t}(x), \mathbf{t}(\gamma^{-1} gy) \rangle,$$

and so

$$\begin{aligned}
& \int_{\mathcal{X}} \int_G h(\kappa \langle \mathbf{t}(\gamma x), \mathbf{t}(gy) \rangle) d\mu(g) d\nu_{\mathcal{X}}(y) \\
= & \int_G \int_{\mathcal{X}} h(\kappa \langle \mathbf{t}(x), \mathbf{t}(\gamma^{-1} gy) \rangle) d\nu_{\mathcal{X}}(y) d\mu(g) \\
= & \int_G \int_{\mathcal{X}} h(\kappa \langle \mathbf{t}(x), \mathbf{t}(z) \rangle) d\nu_{\mathcal{X}}(z) d\mu(g) \\
= & \int_{\mathcal{X}} h(\kappa \langle \mathbf{t}(x), \mathbf{t}(y) \rangle) d\nu_{\mathcal{X}}(y).
\end{aligned}$$

As this does not depend on γ, the marginal density of x is constant. Similarly, the marginal density of y is constant, and so (5) is a copula. A simple extension of this argument shows that (6) is a copula. $\square$

For $\mathcal{X} = S^{p-1}$ and μ a point distribution on $SO(p)$, various subclasses of the densities (5) have been considered elsewhere: (a) taking $\mathbf{t}(\mathbf{x}) = \mathbf{x}$ and $\kappa = 1$ gives the densities of Example 1 of [25]; (b) taking $\mathbf{t}(\mathbf{x}) = \mathbf{x}$ and $h = \exp$ gives a special case of the bivariate densities considered in Section 2.4 of [20] and Section 11.4 of [21], the normalising constant is $c(\kappa) = (\kappa/2)^{p/2-1}\left\{\Gamma(p/2)I_{p/2-1}(\kappa)\right\}^{-1}$, the conditional distribution of $\mathbf{y}$ given $\mathbf{x}$ is von Mises–Fisher, and the correlation coefficient of [16] is $pA_p(\kappa)^2$, where $A_p(\kappa) = I_{p/2}(\kappa)/I_{p/2-1}(\kappa)$; (c) for $\mathbf{t}(\mathbf{x}) = \mathbf{x}\mathbf{x}^{\mathrm{T}} - p^{-1}\mathbf{I}_p$ and $h = \exp$, the normalising constant is $c(\kappa) = M(1/2, p/2, \kappa)^{-1}$, where $M(1/2, p/2, \kappa)$ is a Kummer function, and the conditional distribution of $\mathbf{y}$ given $\mathbf{x}$ is Watson. In the cases $p = 2, 3$ and $\mathbf{t}(\mathbf{x}) = \mathbf{x}$ or $\mathbf{t}(\mathbf{x}) = \mathbf{x}\mathbf{x}^{\mathrm{T}} - p^{-1}\mathbf{I}_p$ (with $\mathbf{I}_p$ denoting the $p \times p$ identity matrix), the corresponding stationary Markov processes were given in models 2.1, 2.4, 3.1 and 3.4 of [1].

4.4. Goodness of fit

Let $f(\cdot,\cdot;\theta)$ (with $\theta \in \Theta$) be a family of copulae on $\mathcal{X}\times\mathcal{Y}$, where $\mathcal{X}$ and $\mathcal{Y}$ are compact Riemannian manifolds. Given observations $(\mathbf{x}_1, \mathbf{y}_1), \ldots, (\mathbf{x}_n, \mathbf{y}_n)$ on $\mathcal{X}\times\mathcal{Y}$ and a corresponding fitted copula $f(\cdot,\cdot;\hat{\theta})$ in this family, the quality of the fit can be assessed using the following version of the Sobolev goodness-of-fit tests of [14]. It is based on the fact that, for distributions on $\mathcal{X}\times\mathcal{Y}$ with uniform marginals, uniformity is equivalent to independence. This implies that, for measuing the fit of a copula, the weighted tests of uniformity used in the goodness-of-fit tests of [14] can be replaced by weighted versions of the tests of independence considered in [18]. Thus, if $\mathbf{t} : \mathcal{X} \to L^2(\mathcal{X})$ and $\mathbf{u} : \mathcal{Y} \to L^2(\mathcal{Y})$ are functions of the form (4), an appropriate statistic for measuring goodness of fit is

$$T = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \langle \mathbf{t}(\mathbf{x}_i), \mathbf{t}(\mathbf{x}_j)\rangle \langle \mathbf{u}(\mathbf{y}_i), \mathbf{u}(\mathbf{y}_j)\rangle, \tag{7}$$

where

$$w_i = f(\mathbf{x}_i, \mathbf{y}_i; \hat{\theta})^{-1} \qquad i = 1, \ldots, n.$$

Large values of T indicate poor fit. Significance of T can be assessed using simulation from the fitted copula.

For certain concentrated copulae of the form (5) or (6), there is a quick graphical method of assessing goodness of fit. If $\mathcal{X}$ is a d-dimensional homogeneous manifold and the copula has density $f(x, y; \kappa, g) = c(\kappa) \exp\{\kappa\langle \mathbf{t}(x), \mathbf{t}(gy)\rangle\}$ then it follows from standard high-concentration asymptotics that a plot of $2\hat{\kappa}\{M - \langle \mathbf{t}(x_1), \mathbf{t}(\hat{g}y_1)\rangle\}, \dots, 2\hat{\kappa}\{M - \langle \mathbf{t}(x_n), \mathbf{t}(\hat{g}y_n)\rangle\}$ (where $M = \langle \mathbf{t}(x), \mathbf{t}(x)\rangle$ and $\hat{g}$ is the maximum likelihood estimate of g) against the $(i-1/2)/n$ quantiles of χ^2_d $(i = 1, \dots, n)$ will be close to the straight line of slope 1 through the origin, provided that $\hat{\kappa}$ is large. Similarly, if $\mathcal{X}$ is a d-dimensional Lie group and the copula has density $f(x, y; \kappa, g_1, g_2) = c(\kappa) \exp\left(\kappa\langle \mathbf{t}(x), \mathbf{t}(g_1 y g_2^{-1})\rangle\right)$ then a plot of $2\hat{\kappa}\left\{M - \langle \mathbf{t}(x_1), \mathbf{t}(\hat{g}_1 y_1 \hat{g}_2^{-1})\rangle\right\}, \dots, 2\hat{\kappa}\left\{M - \langle \mathbf{t}(x_n), \mathbf{t}(\hat{g}_1 y_n \hat{g}_2^{-1})\rangle\right\}$ (where $\hat{g}_1$ and $\hat{g}_2$ are the maximum likelihood estimates of g_1 and g_2) against the $(i-1/2)/n$ quantiles of χ^2_d $(i = 1, \dots, n)$ will be close to the straight line of slope 1 through the origin, provided that $\hat{\kappa}$ is large.

5. Examples

5.1. Magnetic remanence data on $S^2 \times S^2$

Data set B8 of [5] consists of 62 pairs $(\mathbf{x}_1, \mathbf{y}_1), \dots, (\mathbf{x}_n, \mathbf{y}_n)$ of directions of magnetic remanence after partial demagnetisation; $\mathbf{x}_i$ and $\mathbf{y}_i$ are the directions recorded for the ith specimen at 200° C and 350° C, respectively. It is reasonable to model the marginal distributions of $\mathbf{x}$ and $\mathbf{y}$ by Fisher distributions. The maximum likelihood estimates of the mean directions $\boldsymbol{\mu}_X$ and $\boldsymbol{\mu}_Y$ and the concentrations κ_X and κ_Y are $\hat{\boldsymbol{\mu}}_X = (0.210, 0.104, 0.972)^{\mathrm{T}}$, $\hat{\boldsymbol{\mu}}_Y = (0.210, 0.135, 0.968)^{\mathrm{T}}$ and $\hat{\kappa}_X = 76.1$, $\hat{\kappa}_Y = 84.4$. Adequacy of the fit can be assessed by applying tests of unformity to the transformed observations $\hat{\phi}_{\mathcal{X}}(\mathbf{x}_1), \dots, \hat{\phi}_{\mathcal{X}}(\mathbf{x}_n)$ and $\hat{\phi}_{\mathcal{Y}}(\mathbf{y}_1), \dots, \hat{\phi}_{\mathcal{Y}}(\mathbf{y}_n)$, where $\hat{\phi}_{\mathcal{X}}$ and $\hat{\phi}_{\mathcal{Y}}$ are the geodesic-preserving homeomorphisms of S^2 that transform the fitted marginal distributions to uniformity and preserve the sample mean directions; see Example 1 of Section 2. The P-values of the Rayleigh test of uniformity applied to $\hat{\phi}_{\mathcal{X}}(\mathbf{x}_1), \dots, \hat{\phi}_{\mathcal{X}}(\mathbf{x}_n)$ and $\hat{\phi}_{\mathcal{Y}}(\mathbf{y}_1), \dots, \hat{\phi}_{\mathcal{Y}}(\mathbf{y}_n)$ are 0.99 and 0.90, respectively, so that uniformity of the transformed fitted Fisher distributions can certainly be accepted, i.e. the Fisher distributions fit $\mathbf{x}_1, \dots, \mathbf{x}_n$ and $\mathbf{y}_1, \dots, \mathbf{y}_n$ well.

Fitting the Sobolev copula with density

$$f(\mathbf{x}, \mathbf{y}; \kappa, \mathbf{U}) = (\kappa / \sinh \kappa) \exp\{\kappa \mathbf{x}^{\mathrm{T}} \mathbf{U}^{\mathrm{T}} \mathbf{y}\} \tag{8}$$

to the transformed data $(\hat{\phi}_{\mathcal{X}}(\mathbf{x}_1), \hat{\phi}_{\mathcal{Y}}(\mathbf{y}_1)), \ldots, (\hat{\phi}_{\mathcal{X}}(\mathbf{x}_n), \hat{\phi}_{\mathcal{Y}}(\mathbf{y}_n))$ gives maximum likelihood estimates

$$\hat{\kappa} = 6.42, \qquad \widehat{\mathbf{U}} = \begin{pmatrix} 0.997 & -0.074 & 0.013 \\ 0.074 & 0.996 & -0.042 \\ -0.010 & -0.043 & 0.999 \end{pmatrix}.$$

The fit of the copula can be assessed by means of the plot described at the end of Section 4.4, taking $\mathbf{t}(\mathbf{x}) = \mathbf{x}$, and so plotting $2\hat{\kappa}\left\{1 - \hat{\phi}_{\mathcal{X}}(\mathbf{x}_1)^{\mathrm{T}}\widehat{\mathbf{U}}^{\mathrm{T}}\hat{\phi}_{\mathcal{Y}}(\mathbf{y}_1)\right\}, \ldots, 2\hat{\kappa}\left\{1 - \hat{\phi}_{\mathcal{X}}(\mathbf{x}_n)^{\mathrm{T}}\widehat{\mathbf{U}}^{\mathrm{T}}\hat{\phi}_{\mathcal{Y}}(\mathbf{y}_n)\right\}$ against quantiles of χ^2_2. For this copula, the plot is the same as a colatitude plot (see e.g. Section 10.2 of [21]) based on the angles between the $\widehat{\mathbf{U}}\hat{\phi}_{\mathcal{X}}(\mathbf{x}_i)$ and $\hat{\phi}_{\mathcal{Y}}(\mathbf{y}_i)$. For this data set, the plot is far from linear and the corresponding Kolmogorov–Smirnov test is significant at the 1 % level. Thus the copula fits the transformed data very poorly.

5.2. Vectorcardiogram data on $SO(3) \times SO(3)$

The data set described in [4] consists of pairs of rotations representing the orientations of vectorcardiograms of 98 children. The portion of this data set obtained from boys aged 11–19 gives observations $(\mathbf{X}_1, \mathbf{Y}_1), \ldots, (\mathbf{X}_n, \mathbf{Y}_n)$ on $SO(3) \times SO(3)$, where $n = 28$ and rotations $\mathbf{X}_i$ and $\mathbf{Y}_i$ represent the orientations of vectorcardiograms for the ith subject obtained using the Frank lead system and the McFee lead system, respectively, for placement of electrical leads. Prentice [24] fitted the regression model $\mathbf{Y} = \mathbf{U}_1^{\mathrm{T}}\mathbf{X}\mathbf{U}_2$ with $\mathbf{U}_1$ and $\mathbf{U}_2$ in $SO(3)$ separately to the portion of the data obtained from the 56 boys and to that obtained from the 42 girls. Here, the data set on the 28 boys aged 11–19 will be described by the bivariate model with marginal symmetric matrix Fisher distributions and a copula with regression function $\mathbf{Y} = \mathbf{U}_1^{\mathrm{T}}\mathbf{X}\mathbf{U}_2$.

Perhaps the simplest non-uniform models for the marginal distributions of $\mathbf{X}$ and $\mathbf{Y}$ are the matrix Fisher distributions with densities of the form $e^{\kappa_X} M(1/2, 2, 4\kappa_X)^{-1} \exp\{\kappa_X \mathrm{tr}\,(\mathbf{X}^{\mathrm{T}}\mathbf{M}_X)\}$ and $e^{\kappa_Y} M(1/2, 2, 4\kappa_Y)^{-1} \exp\{\kappa_Y \mathrm{tr}\,(\mathbf{Y}^{\mathrm{T}}\mathbf{M}_Y)\}$, respectively, where $\mathbf{M}$ and $\mathbf{M}_Y$ are in $SO(3)$ and $M(1/2, 2, \cdot)$ is a Kummer function. The maximum likelihood estimates of the mean directions $\mathbf{M}_X$ and $\mathbf{M}_Y$ and concentrations κ_X and κ_Y are

$$\widehat{\mathbf{M}}_X = \begin{pmatrix} 0.491 & 0.705 & 0.512 \\ 0.633 & -0.692 & 0.346 \\ 0.598 & 0.154 & -0.786 \end{pmatrix}, \widehat{\mathbf{M}}_Y = \begin{pmatrix} 0.695 & 0.653 & 0.300 \\ 0.639 & -0.753 & 0.160 \\ 0.330 & 0.080 & -0.940 \end{pmatrix}$$

and $\hat{\kappa}_X = 2.76, \hat{\kappa}_Y = 2.76$. The fit of these models to the marginal data can be assessed by applying the Rayleigh test of uniformity to each of the transformed marginal data sets $(\hat{\phi}_{\mathcal{X}}(\mathbf{X}_1), \ldots, (\hat{\phi}_{\mathcal{X}}(\mathbf{X}_n))$ and $(\hat{\phi}_{\mathcal{Y}}(\mathbf{Y}_1), \hat{\phi}_{\mathcal{Y}}(\mathbf{Y}_n))$, where $\hat{\phi}_{\mathcal{X}}$ and $\hat{\phi}_{\mathcal{Y}}$ denote the orientation-preserving homeomorphisms of $SO(3)$ that transform the fitted marginal distributions to uniformity and preserve geodesics through $\widehat{\mathbf{M}}_X$ and $\widehat{\mathbf{M}}_Y$; see Example 2 of Section 2. Comparison of the Rayleigh statistics with their asymptotic distribution (which is justified by [13, Table 1]) yields P-values of 0.001, so the fitted Fisher distributions fit the marginal data very poorly. The Bingham test on $\mathbb{R}P^2$ of uniformity of the rotation axes in each marginal data set gives P-values of 0.62 and 0.06, so that uniformity is acceptable (marginally so for the data on Y). This suggests that a better transformation to marginal normality might be obtained by taking $(\widehat{\mathbf{M}}_X, \widehat{\mathbf{M}}_Y)$ as the base-point of $SO(3) \times SO(3)$ and using the uniform scores $(\tilde{X}_1, \tilde{Y}_1), \ldots, (\tilde{X}_n, \tilde{Y}_n)$ (described in Section 2.2) instead of the transformed data $(\hat{\phi}_{\mathcal{X}}(\mathbf{X}_1), \hat{\phi}_{\mathcal{Y}}(\mathbf{Y}_1)), \ldots, (\hat{\phi}_{\mathcal{X}}(\mathbf{X}_n), \hat{\phi}_{\mathcal{Y}}(\mathbf{Y}_n))$.

Fitting the copula with density

$$f(\mathbf{X}, \mathbf{Y}; \kappa, \mathbf{U}_1, \mathbf{U}_2) = e^{\kappa} M(1/2, 2, 4\kappa)^{-1} \exp\left\{\kappa \mathrm{tr}(\mathbf{X}^{\mathrm{T}} \mathbf{U}_1 \mathbf{Y} \mathbf{U}_2^{\mathrm{T}})\right\}$$

(which has the form (6) with $\mathbf{t}(\mathbf{X}) = \mathbf{u}(\mathbf{X}) = \mathbf{X}$) to the uniform scores $(\tilde{X}_1, \tilde{Y}_1), \ldots, (\tilde{X}_n, \tilde{Y}_n)$ gives maximum likelihood estimates

$$\widehat{\mathbf{U}}_1 = \begin{pmatrix} 0.980 & 0.198 & -0.037 \\ -0.201 & 0.950 & -0.241 \\ -0.013 & 0.244 & 0.970 \end{pmatrix}, \widehat{\mathbf{U}}_2 = \begin{pmatrix} 0.988 & 0.153 & 0.024 \\ -0.147 & 0.974 & -0.172 \\ -0.050 & 0.166 & 0.985 \end{pmatrix}$$

and $\hat{\kappa} = 0.80$. A weighted Sobolev goodness-of-fit test based on (7) with $\mathbf{t}(\mathbf{X}) = \mathbf{u}(\mathbf{X}) = \mathbf{X}$ has P-value 0.16, indicating that the fit of the copula to the uniform scores $(\tilde{X}_1, \tilde{Y}_1), \ldots, (\tilde{X}_n, \tilde{Y}_n)$ is acceptable.

References

[1] L. Accardi, J. Cabrera, G.S. Watson, Some stationary Markov processes in discrete time for unit vectors, Metron 45 (1987) 115–133.

[2] W.M. Boothby, An Introduction to Differentiable Manifolds and Riemannian Geometry. 2nd. ed., Academic Press, San Diego, 2003.

[3] W.F. Darsow, B. Nguyen, E.T. Olsen, Copulas and Markov processes, Illinois J. Math. 36 (1992) 600–642.

[4] T.D. Downs, J. Liebman, W. Mackay, Statistical methods for vectorcardiogram orientations, Proc. XIth International Vectorcardiogram Symposium (I. Hoffman, R. I. Hamby and E. Glassman, Eds.) 216–222. North-Holland, Amsterdam, 1971. Data available at `http://www.mcs.st-andrews.ac.uk/~pej/vcg.html`.

[5] N.I. Fisher, T. Lewis, B.J.J. Embleton, Statistical Analysis of Spherical Data, Cambridge University Press, Cambridge, 1987.

[6] S. Gallot, D. Hulin, J. Lafontaine, Riemannian Geometry. 2nd. ed., Springer-Verlag, Berlin, 1993.

[7] E. Giné M., Invariant tests for uniformity on compact Riemannian manifolds based on Sobolev norms, Ann. Statist. 3 (1975) 1243–1266.

[8] V. Isham, A Markov construction for a multidimensional point process, J. Appl. Prob. 14 (1977) 507–515.

[9] H. Joe, Multivariate Models and Dependence Concepts, Chapman & Hall, London, 1997.

[10] R.A. Johnson, T.E. Wehrly, Measures and models for angular correlation and angular-linear correlation, J. Roy. Statist. Soc. Ser. B. 39 (1977) 222–229.

[11] R.A. Johnson, T.E. Wehrly, Some angular-linear distributions and related regression models, J. Amer. Statist. Assoc. 73 (1978) 602–606.

[12] M.C. Jones, A. Pewsey, S. Kato, On a class of circulas: copulas for circular distributions, Ann. Inst. Statist. Math. DOI 10.1007/s10463-014-0493-6.

[13] P.E. Jupp, Modifications of the Rayleigh and Bingham tests for uniformity of directions, J. Mult. Anal. 77 (2001) 1–20.

[14] P.E. Jupp, Sobolev tests of goodness of fit of distributions on compact Riemannian manifolds, Ann. Statist. 33 (2005) 2957–2966.

[15] P.E. Jupp, Data-driven Sobolev tests of uniformity on compact Riemannian manifolds, Ann. Statist. 36 (2008) 1246–1260.

[16] P.E. Jupp, K.V. Mardia, A general correlation coefficient for directional data and related regression problems, Biometrika 67 (1980) 163–173.

[17] P.E. Jupp, B.D. Spurr, Sobolev tests for symmetry of directional data, Ann. Statist. 11 (1983) 1225–1231.

[18] P.E. Jupp, B.D. Spurr, Sobolev tests for independence of directions, Ann. Statist. 13 (1985) 1140–1155.

[19] J.T. Kent, The infinite divisibility of the von Mises–Fisher distribution for all values of the parameter in all dimensions, Proc. London Math. Soc. 35 (1977) 359–384.

[20] K.V. Mardia, Statistics of directional data, J. Roy. Statist. Soc. Ser. B. 37 (1975) 349–393.

[21] K.V. Mardia, P.E. Jupp, Directional Statistics, Wiley, Chichester, 2000.

[22] J. Moser, On the volume elements on a manifold, Trans. Amer. Math. Soc. 120 (1965) 286–294.

[23] R.N. Nelsen, An Introduction to Copulas, 2nd ed., Springer, New York, 2006.

[24] M.J. Prentice, Spherical regression on matched pairs of orientation statistics, J. Roy. Statist. Soc. B 51 (1989) 241–248.

[25] L.-P. Rivest, Symmetric distributions for dependent unit vectors, Ann. Statist. 12 (1984) 1050–1057.

[26] M. Scarsini, Copulae of probability measures on product spaces, J. Mult. Anal. 31 (1989) 201–219.

[27] A. Sklar, Fonctions de répartition à n dimensions et leur marges, Publ. Inst. Statist. Univ. Paris 8 (1959) 229–231.

[28] T.E. Wehrly, R.A. Johnson, Bivariate models for dependence of angular observations and a related Markov process, Biometrika 67 (1980) 255–256.

[29] J.A. Wellner, Permutation tests for directional data, Ann. Statist. 7 (1979) 929–943.