METMA V International Workshop on Spatio-Temporal Modelling
Santiago de Compostela
30 Jun - 2 Jul 2010

Analyzing Circular Time Series by Using Markov Models Based on Nonnegative Trigonometric Sums

J.J. Fernández-Durán[1,*] and M.M. Gregorio-Domínguez[2]

[1] *Departament of Statistics, División Académica de Actuaría, Estadística y Matemáticas, Instituto Tecnológico Autónomo de México. Río Hondo No. 1, Col. Progreso Tizapán, C.P. 01080, México D.F., México, e-mail: jfdez@itam.mx*
[2] *Departament of Actuarial Science, División Académica de Actuaría, Estadística y Matemáticas, Instituto Tecnológico Autónomo de México. Río Hondo No. 1, Col. Progreso Tizapán, C.P. 01080, México D.F., México, e-mail: mercedes@itam.mx*
[*] *Corresponding author*

Abstract. *In [2] a new family of univariate circular distributions based on nonnegative trigonometric (Fourier) sums (series) was developed. This family was extended to the multivariate case in [4] by using multiple nonnegative trigonometric sums to model the joint distribution of a vector of angular (circular) random variables. In this work, we use the univariate and multivariate family of circular distributions based on nonnegative trigonometric sums to model time series of circular data by using Markov models. Practical examples of circular time series include the hourly wind directions at different monitoring stations, the hourly recorded directions taken by an animal and the hourly recorded directions of a hurricane, among others.*

Keywords. *Fourier Series; Non-negative Fourier series.*

1 Univariate and Multivariate Nonnegative Trigonometric Sums (NNTS) Densities

The probability density function, $f(\theta;\underline{c})$, of a circular random variable $\Theta \in (0,2\pi]$ must be nonnegative and periodic ($f(\theta+2k\pi;\underline{c}) = f(\theta;\underline{c})$) for any integer k where $\underline{c}$ is the vector of parameters (see [5], [7], [8] and [9]).

[1] first identified the conditions that must be satisfied by the coefficients of a nonnegative trigono-

metric sum. Using Fejér's findings, [2] defined a family of circular distributions by

$$f(\theta;\underline{c},M) = \left|\left|\sum_{k=0}^{M} c_k e^{ik\theta}\right|\right|^2 = \sum_{k=0}^{M}\sum_{l=0}^{M} c_k\bar{c}_l e^{i(k-l)\theta} = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{k=1}^{M} a_k cos(k\theta) + b_k sin(k\theta) \tag{1}$$

where $\sum_{k=0}^{M}||c_k||^2 = \frac{1}{2\pi}$ and $a_k - ib_k = 2\sum_{\nu=0}^{n-k} c_{\nu+k}\bar{c}_\nu$ for complex numbers $c_l = c_{rl} + ic_{cl}$ for $k = 1,2,...,M$ and $l = 0,1,...,M$. Note that the uniform circular distribution can be seen as a particular case of the family by setting $a_k = b_k = 0$ for $k = 1,...,M$ or, equivalently, $M = 0$. Given the restriction $\sum_{k=0}^{M}||c_k||^2 = \frac{1}{2\pi}$, c_0 is a positive real number $c_0 = \sqrt{\frac{1}{2\pi} - \sum_{k=1}^{M}||c_k||^2}$. In this case, the vector of complex parameters, $\underline{c}$, has a dimension of $2M$+2 with $2M$ free parameters. Also, the order of the trigonometric sum, M, which also corresponds to the maximum number of modes of the density, is considered as an additional parameter. This family of circular distributions has the advantage of being able to fit datasets that present multimodality and/or skewness because the density function can be expressed as a mixture of multimodal circular distributions. The interpretation of the $\underline{c}$ parameters can be done by expressing the c parameters as $c_k = \rho_k e^{-i\phi_k}$ for $k = 0,\ldots,M$. Then, the angles of the c parameters, ϕ_k for $k = 0,\ldots,M$, are location parameters and the norms of the c parameters, ρ_k for $k = 0,\ldots,M$, are scale parameters that reflect the importance of each of the terms in the sum that defines the circular density in Equation (1) because $\sum_{k=0}^{M}\rho_k^2 = \frac{1}{2\pi}$.

Let $\underline{\theta} = (\theta_1,\theta_2,\ldots,\theta_R)$ be a vector of angular random variables, i.e., $\theta_i \in (0,2\pi]$ for $i = 1,\ldots,R$ and let $f(\underline{\theta};\underline{c})$ be its probability density function. Note that $f(\underline{\theta};\underline{c})$ must have a period of 2π in all its arguments.

Now let us extend the results of [2] and [3] to multivariate random vectors of angular random variables. We consider multiple trigonometric (Fourier) sums (series). Let the squared norm of such a sum define the density function $f(\underline{\theta};\underline{c},\underline{M})$ as

$$\begin{aligned} f(\underline{\theta};\underline{c},\underline{M}) &= \left|\left|\sum_{k_1=0}^{M_1}\sum_{k_2=0}^{M_2}\cdots\sum_{k_R=0}^{M_R} c_{k_1k_2\ldots k_R} e^{i(k_1\theta_1+k_2\theta_2+\ldots+k_R\theta_R)}\right|\right|^2 \\ &= \sum_{k_1=0}^{M_1}\sum_{m_1=0}^{M_1}\sum_{k_2=0}^{M_2}\sum_{m_2=0}^{M_2}\cdots\sum_{k_R=0}^{M_R}\sum_{m_R=0}^{M_R} c_{k_1k_2\ldots k_R}\bar{c}_{m_1m_2\ldots m_R} e^{i\sum_{j=1}^{R}(k_j-m_j)\theta_j} \end{aligned} \tag{2}$$

where the vector of parameters $(\underline{c},\underline{M})$ contains the complex numbers $c_{k_1k_2\ldots k_R}$ for $k_j = 0,\ldots,M_j$ for $j = 1,\ldots,R$ and $\underline{M} = (M_1,M_2,\ldots,M_R)$ are additional parameters and $\bar{c}_{m_1m_2\ldots m_R}$ refers to the complex conjugate of $c_{m_1m_2\ldots m_R}$. To ensure that $f(\underline{\theta};\underline{c},\underline{M})$ integrates to 1 the following restriction on the $\underline{c}$ parameter must be imposed

$$\sum_{k_1=0}^{M_1}\sum_{k_2=0}^{M_2}\cdots\sum_{k_R=0}^{M_R}||c_{k_1k_2\ldots k_R}||^2 = \frac{1}{(2\pi)^R}. \tag{3}$$

Note that the special case $M_1 = M_2 = \ldots = M_R = 0$ corresponds to a uniform distribution on the R-dimensional torus. The components of the vector $\underline{M}$ that are equal to zero correspond to marginal uniform distributions of the corresponding components of the $\underline{\theta}$ random vector.

The joint trigonometric moment of order $(n_1,n_2,\ldots,n_R)$ for integers n_j for $j = 1,\ldots,R$ is given by

$$E\left(e^{i\sum_{j=1}^{R} n_j\theta_j}\right) = (2\pi)^R \sum_{k_1-m_1=n_1}^{M_1} \sum_{k_2-m_2=n_2}^{M_2} \cdots \sum_{k_R-m_R=n_R}^{M_R} c_{k_1k_2\ldots k_R}\bar{c}_{m_1m_2\ldots m_R}$$
$$= (2\pi)^R \sum_{m_1=0}^{M_1-n_1} \sum_{m_2=0}^{M_2-n_2} \cdots \sum_{m_R=0}^{M_R-n_R} c_{m_1+n_1,m_2+n_2,\ldots,m_R+n_R}\bar{c}_{m_1m_2\ldots m_R}. \quad (4)$$

The univariate marginal distributions, $f(\theta_n;\underline{c}^{(n)},M_n)$ are given by

$$f(\theta_n;\underline{c}^{(n)},M_n) = (2\pi)^{R-1}\sum_{k_n}^{M_n}\sum_{m_n}^{M_n} c^{(n)}_{k_nm_n} e^{i(k_n-m_n)\theta_n} \quad (5)$$

where $c^{(n)}_{k_nm_n}$ is equal to

$$\sum_{k_1=0}^{M_1}\sum_{k_2=0}^{M_2}\cdots\sum_{k_{n-1}=0}^{M_{n-1}}\sum_{k_{n+1}=0}^{M_{n+1}}\cdots\sum_{k_R=0}^{M_R} c_{k_1k_2\ldots k_{n-1}k_nk_{n+1}\ldots k_R}\bar{c}_{k_1k_2\ldots k_{n-1}m_nk_{n+1}\ldots k_R}. \quad (6)$$

Note that the univariate marginal distributions are members of the multivariate nonnegative trigonometric sums (MNNTS) family. This finding also applies to the marginal distributions of any number of components of the original vector of angular random variables having an MNNTS distribution. Conditional densities are obtained from the joint distributions and are also members of the MNNTS family. Note that the joint, marginal and conditional trigonometric moments are obtained as functions of the $\underline{c}$ parameters.

2 Circular Time Series Markov Models Based on NNTS Densities

For previous works in the statistical analysis of time series of circular data, the interested reader can consult [6] and the references therein. Now, since the (M)NNTS $\underline{c}$ parameter space corresponds to the surface of a hypersphere, it is possible to construct Markov models for time series of univariate or multivariate circular data: Let $\theta_1,\theta_2,\ldots,\theta_T$ be a time series of (univariate or multivariate) circular data. We assume that θ_t is distributed as an (M)NNTS density with parameters $(\underline{c}_t,\underline{M})$ for $t=1,2,\ldots,T$. The equation defining the Markovian evolution of the $\underline{c}$ parameter is given as follows:

$$\underline{c}_t = E\underline{c}_{t-1} + v_t \quad (7)$$

where E is a complex unitary matrix and v_t is a random error on the surface of the corresponding hypersphere (for example, von Mises-Fisher distributed). The matrix E should be complex unitary to maintain the norm constraint of $\underline{c}_t$:

$$||E\underline{c}_{t-1}||^2 = \underline{c}'_{t-1}E'E\underline{c}_{t-1} = \underline{c}'_{t-1}\underline{c}_{t-1} \quad (8)$$

since $E'E=I$. We applied the proposed methodology to two real datasets. The first one consists of the circular time series of directions of hurricanes and the second one is on levels of certain pollutants and wind directions in Mexico City. The method of maximum likelihood was used to estimate the parameters, E and $\underline{M}$, of the (M)NNTS Markov time series model.

Acknowledgments. The authors thank the Asociación Mexicana de Cultura for its support.

References

[1] Fejér, L. (1915) Über trigonometrische Polynome. *Journal fur die Reine und Angewandte Mathematik.*, **146**, pp. 53-82.

[2] Fernández-Durán, J.J. (2004) Circular Distributions Based on Nonnegative Trigonometric Sums. *Biometrics*, **60**, pp. 499-503.

[3] Fernández-Durán, J.J. (2007) Models for Circular-Linear and Circular-Circular Data Constructed from Circular Distributions Based on Nonnegative Trigonometric Sums. *Biometrics*, **63**, pp.579-585.

[4] Fernández-Durán, J.J. and Gregorio-Domínguez, M.M (2009) Multivariate Angular Distributions Based on Multiple Nonnegative Trigonometric Sums. *Working Paper, Department of Statistics, ITAM*, **DE-C09.1**.

[5] Fisher, N.I. (1993) *Statistical Analysis of Circular Data*. Cambridge University Press. Cambridge.

[6] Holzmann, H., Munk, A., Suxter, M. and Zucchini, W. (2006) Hidden Markov Models for Circular and Linear-Circular Time Series. *Environmental and Ecological Statistics*, **13**, pp. 325-347.

[7] Jammalamadaka, S.R. and SenGupta, A. (2001) *Topics in Circular Statistics*. World Scientific Publishing, Co., Singapore.

[8] Mardia, K.V. and Jupp, P.E. (2000) *Directional Statistics*. John Wiley and Sons. Great Britain.

[9] Upton, G.J.G. and Fingleton, B. (1989) *Spatial Data Analysis by Example Vol. 2 (Categorical and Directional Data)*. John Wiley and Sons. Great Britain.