Journal of Mathematical Psychology 80 (2017) 4–14

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

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

Bayesian estimation and hypothesis tests for a circular Generalized Linear Model

Kees Mulder [a,*], Irene Klugkist [a,b]

[a] Department of Methodology and Statistics, Utrecht University, The Netherlands
[b] Research Methodology, Measurement and Data Analysis, Behavioural Sciences, University of Twente, Enschede, The Netherlands

HIGHLIGHTS

- A Bayesian analysis of circular data using a GLM-type model based on the von Mises distribution is proposed.
- A weakly informative prior solves issues that are common for this model in a frequentist setting.
- Hypothesis tests are developed for both equality and inequality constrained hypotheses.
- The model is shown to work well and provide valuable insight for psychological research.

ARTICLE INFO

Article history:
Received 16 September 2016
Received in revised form 15 May 2017
Available online 23 August 2017

Keywords:
Circular data
MCMC
Bayes factor
Savage–Dickey density ratio

ABSTRACT

Motivated by a study from cognitive psychology, we develop a Generalized Linear Model for circular data within the Bayesian framework, using the von Mises distribution. Although circular data arise in a wide variety of scientific fields, the number of methods for their analysis is limited. Our model allows inclusion of both continuous and categorical covariates. In a frequentist setting, this type of model is plagued by the likelihood surface of its regression coefficients, which is not logarithmically concave. In a Bayesian context, a weakly informative prior solves this issue, while for other parameters noninformative priors are available. In addition to an MCMC sampling algorithm, we develop Bayesian hypothesis tests based on the Bayes factor for both equality and inequality constrained hypotheses. In a simulation study, it can be seen that our method performs well. The analyses are available in the package CircGLMBayes. Finally, we apply this model to a dataset from experimental psychology, and show that it provides valuable insight for applied researchers. Extensions to dependent observations are within reach by means of the multivariate von Mises distribution.

1. Introduction

Circular data are measured in angles or directions, and are frequently encountered in scientific fields as diverse as life sciences (Mardia, 2011), behavioral biology (Bulbert, Page, & Bernal, 2015), cognitive psychology (Kaas & Van Mier, 2006), bioinformatics (Mardia, Hughes, Taylor, & Singh, 2008), political science (Gill & Hangartner, 2010) and environmental sciences (Arnold & SenGupta, 2006; Lagona, 2016; Lagona, Picone, Maruotti, & Cosoli, 2015). In psychology, circular data occur often in motor behavior research (Baayen, Klugkist, & Mechsner, 2012; Mechsner, Kerzel, Knoblich, & Prinz, 2001; Mechsner, Stenneken, Cole, Aschersleben, & Prinz, 2007; Postma, Zuidhoek, Noordzij, & Kappers, 2008), as well as in the application of circumplex models (Gurtman, 2009; Gurtman & Pincus, 2003; Leary, 1957). Circular data differ from linear data in the sense that circular data are measured in a periodical sample space. For example, an angle of 1° is quite close to an angle 359°, although linear intuition suggests otherwise.

Therefore, linear models may not properly describe the process that has generated the circular data of interest. Circular data analysis has been developed to deal with this, although attention to this type of analysis has been limited. Only slightly more than a handful of in-depth books on circular data analysis have been published (Fisher, 1995; Jammalamadaka & Sengupta, 2001; Mardia & Jupp, 1999; Pewsey, Neuhäuser, & Ruxton, 2013), and in general, statistical methods for circular data are somewhat limited.

Here, attention is turned to analysis of datasets with a circular outcome, predicted by covariates that can be continuous (linear) or categorical. This leads to a structure similar to the Generalized Linear Model (GLM), which has both multiple regression and ANCOVA as special cases.

* Corresponding author.
E-mail address: k.t.mulder@uu.nl (K. Mulder).

http://dx.doi.org/10.1016/j.jmp.2017.07.001
0022-2496/

Three main approaches to circular data analysis might be distinguished. First, the intrinsic approach employs distributions directly defined on the circle (Artes, 2008; Fisher & Lee, 1992). Second, the wrapping approach 'wraps' a univariate distribution around the circle by taking the modulus of data on the real line (Coles, 1998; Ferrari, unpublished). Third, the embedding approach projects points from a bivariate distribution to the circle (Hernandez-Stumpfhauser, Breidt, van der Woerd, et al., 2015; Maruotti, 2016; Nuñez-Antonio & Gutiérrez-Peña, 2014; Nuñez-Antonio, Gutiérrez-Peña, & Escarela, 2011; Wang & Gelfand, 2014). While the wrapping and embedding approach provide promising avenues of study in their own right, here attention is restricted to the intrinsic approach, as it might provide the most natural analysis of circular data.

Within the intrinsic approach, the circular analogue to the Normal distribution is the von Mises distribution (Von Mises, 1918). This symmetric unimodal distribution is given by

$$\mathcal{M}(\theta \mid \mu, \kappa) = [2\pi I_0(\kappa)]^{-1} \exp\left(\kappa \cos\left[\theta - \mu\right]\right), \tag{1}$$

where $\theta \in (-\pi, \pi)$ represents an angular measurement, $\mu \in (-\pi, \pi)$ represents the mean direction, $\kappa \in \mathbb{R}^+$ is a concentration parameter, and $I_0(\cdot)$ represents the modified Bessel function of the first kind and order zero. Some examples of frequentist methods that employ the von Mises distribution are a circular ANOVA (Watson & Williams, 1956), circular ANCOVA (Artes, 2008) and circular regression (Fisher & Lee, 1992). Here, a Bayesian analysis of such models will be developed.

Early approaches to Markov chain Monte Carlo (MCMC) sampling for the von Mises distribution provide a method for sampling μ when κ is known (Mardia & El-Atoum, 1976) and sampling both parameters for a single group of data (Damien & Walker, 1999). Guttorp and Lockhart (1988) present a conjugate prior for the von Mises model. Recent theoretical work has much improved the efficiency of the sampling of the concentration parameter of the von Mises distribution (Forbes & Mardia, 2015).

Some development has also taken place in the field of semiparametric inference for circular data models, often using Dirichlet process priors (Bhattacharya & Sengupta, 2009; George & Ghosh, 2006; Ghosh, Jammalamadaka, & Tiwari, 2003; McVinish & Mengersen, 2008). In particular, Ghosh et al. (2003) provide Bayes factors for the simple hypothesis test of equality of two means. However, these methods are generally complex, which makes it hard to extend these models, for example to include covariates. Therefore, we will focus on parametric models, with residuals following the von Mises distribution.

A Bayesian circular regression analysis has been developed by Gill and Hangartner (2010), using starting values from a frequentist iterative reweighted least squares (IRLS) algorithm, which is similar to that used by Fisher and Lee (1992). Gill and Hangartner (2010) note that the likelihood function of the regression coefficients from their model is not globally logarithmically concave, which might cause the algorithm to converge to a local maximum. To combat this, Gill and Hangartner (2010) advise careful inspection of the likelihood surface of the regression coefficients. Drawbacks of the approach taken by Gill and Hangartner (2010) are that a prior is not specified, the algorithm is slow, categorical predictors are not treated separately and for larger models it may be unclear whether the regression coefficients have converged to the global maximum.

Recent work has provided a multivariate extension of the von Mises distribution (Mardia et al., 2008; Mardia & Voss, 2014), which offers a promising new way of thinking about circular covariate models. The multivariate von Mises was applied in this context by Lagona (2016) within a Generalized Linear Model (GLM) setting, applying MCMC likelihood approximation as in Geyer and Thompson (1992) to compute maximum likelihood estimates. This approach is not Bayesian, but it is a promising approach because of its flexibility, allowing both the mean and concentration to be dependent on an arbitrary set of covariates, as well as allowing observations to be dependent.

There are three main drawbacks of the circular GLM approach to circular data analysis currently. First, the GLM approach is not free from the lack of concavity as described in Gill and Hangartner (2010), although this has not yet been investigated in detail. Second, the current approach does not have separate parameters for differences in group mean direction, which precludes the popular ANCOVA model to some extent. Third, Bayesian hypothesis tests for this model are not available, which limits its applicability.

The structure of this paper is as follows. The circular data GLM model is developed in a fully Bayesian setting in Section 2. The lack of concavity in the likelihood function will be examined, and suggestions will be formulated on how to deal with this issue. Details on the MCMC sampler are provided in Section 3. Section 4 outlines Bayesian hypothesis tests for this model, both for equality and inequality constrained hypotheses. Then, a simulation study for the method is provided in Section 5. Section 6 provides an application of our method to empirical data from cognitive psychology. Finally, Section 7 provides a short discussion.

2. Bayesian circular GLM

Consider a dataset $\{\theta_i, \boldsymbol{x}_i, \boldsymbol{d}_i\}, (i = 1, \ldots, n)$, where $\theta_i \in [-\pi, \pi)$ is a circular outcome variable, $\boldsymbol{x}_i \in \mathbb{R}^K$ is a column vector of continuous linear covariates which are assumed to be standardized, and $\boldsymbol{d}_i \in \{0, 1\}^J$ is a column vector of dichotomous variables indicating group membership. Assume that each observed angle θ_i is generated independently from a von Mises distribution $\mathcal{M}(\theta_i \mid \mu_i, \kappa)$. Then, μ_i is chosen to be

$$\mu_i = \beta_0 + \boldsymbol{\delta}^T \boldsymbol{d}_i + g(\boldsymbol{\beta}^T \boldsymbol{x}_i), \tag{2}$$

where $\beta_0 \in [-\pi, \pi)$ is an offset parameter which serves as a circular intercept, $\boldsymbol{\delta} \in [-\pi, \pi)^J$ is a column vector of circular group difference parameters, $g(\cdot) : \mathbb{R} \to (-\pi, \pi)$ is a twice differentiable link function, and $\boldsymbol{\beta} \in \mathbb{R}^K$ is a column vector of regression coefficients. Fisher and Lee (1992) and Jammalamadaka and Sengupta (2001) discuss the choice of the link function. A common and natural choice for the link function is $g(x) = 2\tan^{-1}(x)$, which we will focus on here.

This model specification differs from the usual approach to circular regression models, as these generally set $\mu_i = \beta_0 + g(\boldsymbol{\beta}^T \boldsymbol{x}_i)$ (Fisher & Lee, 1992; Gill & Hangartner, 2010; Lagona, 2016). However, we view this model as unsatisfactory when including dichotomous predictors in $\boldsymbol{x}$, which we will illustrate in Fig. 1. Consider a single dichotomous predictor d added to a model with a single continuous predictor x. The dichotomous predictor might be added into the model as $\mu = \beta_0 + g(\beta x + \delta d)$. Adding δ in the link function shifts the location of the prediction line, but also its shape. Therefore, the shape for $d = 0$ is fixed, but for $d = 1$ the shape is dependent on a free parameter, δ. This makes the shape of the prediction line (and therefore the analysis) depend on the arbitrary choice of reference group, which can be seen in Fig. 1a. To solve this, we advocate setting $\mu = \beta_0 + \delta d + g(\beta x)$, the resulting prediction lines of which are shown in Fig. 1b.

A comparable approach is taken in Artes (2008), where a separate intercept is estimated for each group. However, having a separate intercept for each group means that a factorial design with main effects only cannot be specified. In many applications, especially in psychology, this is problematic. The approach here is more flexible in that it allows a researcher to either fit a model with main effects only, to fit a model with specific interactions, or to compare these models. In addition, Artes (2008) also describes a non-parallel case where the regression parameters are estimated separately for each group. This model can be obtained as a special case of the model provided here by including appropriate interaction terms in the model.