

Research Report

KSTS/RR-06/006

May 24, 2006

A circular-circular regression model

by

Shogo Kato, Kunio Shimizu

Grace S. Shieh

Shogo Kato Keio University Kunio Shimizu Keio University Grace S. Shieh Academia Sinica
--

Department of Mathematics
Faculty of Science and Technology
Keio University

©2006 KSTS

3-14-1 Hiyoshi, Kohoku-ku, Yokohama, 223-8522 Japan

A circular–circular regression model

SHOGO KATO^{1,*}, KUNIO SHIMIZU² AND GRACE S. SHIEH³

¹*School of Fundamental Science and Technology, Keio University*

²*Department of Mathematics, Keio University*

³*Institute of Statistical Science, Academia Sinica*

Abstract

This paper provides a regression model in which both dependent and independent variables are angular. The regression curve is expressed as a form of the Möbius transformation. The angular error is assumed to follow a wrapped Cauchy or, equivalently, circular Cauchy distribution. A bivariate circular distribution is proposed to model our circular regression. Some properties of the regression regression, including estimation and testing procedures, are obtained. The proposed methods are applied to marine biology and wind directions data. A related multiple circular regression model is also introduced.

Key words and phrases: bivariate circular distribution, Möbius transformation, multiple circular regression, wrapped Cauchy distribution.

1 Introduction

Some regression models in which both dependent and independent variables take values on the circle have been proposed in the literature. Rivest (1997) provided a model for predicting the y -direction using a rotation of the “decentred” x -angle, which was applied to the prediction of the direction of earthquake displacement in terms of the direction of steepest descent. Downs and Mardia (2002) proposed a regression model in which the regression curve is expressed as a form of the Möbius transformation or tangent function, with application to data on circadian biological rhythms and wind directions. See Fisher (1993, p.168) for earlier works on circular–circular regression models.

The Möbius transformation is well known as a mapping which carries the unit circle onto itself. One of the earlier works in directional statistics in which the Möbius transformation appeared was given by McCullagh (1996). In this paper he discussed the connection between the standard Cauchy distribution and the wrapped or circular Cauchy distribution via the Möbius transformation. The Möbius transformation was also used in the link functions of regression models by Downs and Mardia (2002) and Downs (2003). Minh and Farnum (2003) induced some probabilistic models on the circle by using a bilinear transformation which maps the real line onto the unit circle and is related to the Möbius transformation in form. Jones (2004) proposed the Möbius distribution on the

*Corresponding author. Tel.: +81-45-566-1641; fax: +81-45-566-1642.
e-mail: shogo@stat.math.keio.ac.jp

disc which is generated by applying the Möbius transformation to the symmetric beta or Pearson type II distribution. A related work, McCullagh (1989), considered the transformation of the distribution on the open interval $(-1, 1)$ via a one-to-one mapping which maps $(-1, 1)$ on to itself. Seshadri (1991) also used a one-to-one mapping, the Möbius transformation on $(-1, 1)$, to investigate the properties of the distribution on $(-1, 1)$.

The wrapped Cauchy distribution was used as a statistical model by Mardia (1972, p.56) and Mardia and Jupp (2000, p.51). Its distributional properties and estimation were investigated by Kent and Tyler (1988) and McCullagh (1996). McCullagh (1996) showed that the wrapped Cauchy distribution is obtained by applying a bilinear transformation to the Cauchy distribution on the real line and is closed under the Möbius transformation. It has the additive property and a central limit theorem holds for this distribution (Kolassa and McCullagh, 1990).

In this paper we propose a new circular-circular regression model and study some properties, including estimation and testing procedures, of this model. Its regression curve is expressed as a form of the Möbius transformation. The angular error is distributed as a wrapped Cauchy distribution.

In Section 2 some properties of the proposed model, including its regression curve and the probability distribution of the angular error, are investigated. In addition, we compare our regression model with some existing models. A bivariate circular distribution, which could be useful for our regression model, is presented in Section 3. Next Section 4 considers parameter estimation, the Fisher information matrix, and a test of independence for the proposed model. In Section 5 our model is applied to marine biology and wind directions data. A circular multiple regression model is introduced in Section 6.

2 Regression Model

2.1 Regression curve

Let u be a variable, which takes values on the unit circle in the complex plane. Suppose β_0 and β_1 are complex parameters with $|\beta_0| = 1$ and $\beta_1 \in \mathbb{C}$. Then the regression curve of the proposed regression model is given by

$$v = \beta_0 \frac{u + \beta_1}{1 + \overline{\beta_1}u}, \quad |u| = 1, \quad (2.1)$$

where the mapping with $|\beta_1| \neq 1$ will be called the Möbius transformation. It is known that the Möbius transformation is a one-to-one mapping which carries the unit circle onto itself.

The Möbius transformation is obtained by a superposition of transformations of the following four types:

- (1) Translation: $z \rightarrow z + b$,
- (2) Rotation: $z \rightarrow az, \quad |a| = 1$,
- (3) Homotheties: $z \rightarrow rz, \quad r > 0$,
- (4) Inversion: $z \rightarrow 1/z$.

For $\beta_1 \neq 0$, the above follow from the identity

$$v = \beta_0 \left(\frac{1}{\beta_1} + \frac{\lambda}{\beta_1 u + 1} \right), \quad \lambda = \beta_1 - \frac{1}{\beta_1}.$$

In (2.1) β_0 is evidently a rotation parameter, but the interpretation of β_1 is more complicated. However, the following properties reveal the function of β_1 in (2.1) for $|\beta_1| < 1$. Assume, without loss of generality, that $\beta_0 = 1$. Then the following results are readily established.

- (a) $|\beta_1| \rightarrow 0 \implies v \rightarrow u$,
- (b) $|\beta_1| \rightarrow 1, u \neq -\beta_1/|\beta_1| \implies v \rightarrow \beta_1/|\beta_1|$,
- (c) $v_j = (u + \beta_{1j})/(1 + \overline{\beta_{1j}}u)$, $\beta_{1j} = r_j e^{i\theta}$, $j = 1, 2$, $r_1 > r_2 \geq 0$, $0 \leq \theta < 2\pi$
 $\implies |\arg(v_1) - \theta| \leq |\arg(v_2) - \theta|$,
- (d) $|\arg(u) - \arg(\beta_1)| \geq |\arg(v) - \arg(\beta_1)|$,
- (e) $u_1 = \beta_1/|\beta_1| \implies v(u_1) = \beta_1/|\beta_1|$,
- (f) $u_2 = -\beta_1/|\beta_1| \implies v(u_2) = -\beta_1/|\beta_1|$,
- (g) $u_1 = \theta\beta_1/|\beta_1|$, $u_2 = \overline{\theta}\beta_1/|\beta_1|$, $|\theta| = 1 \implies v(u_1)\overline{\theta} = \overline{v(u_2)}\theta$.

From these facts, β_1 can be intuitively interpreted as the parameter that attracts the points on the circle toward $\beta_1/|\beta_1|$ with the concentration of points about $\beta_1/|\beta_1|$ increasing as $|\beta_1|$ increases. An exception is the point $u = -\beta_1/|\beta_1|$, which is invariant under the Möbius transformation for any $|\beta_1| < 1$.

The four frames of Figure 1 exhibit the behaviour of (2.1) for selected values of β_1 . Figure 1 (a) is a case in which $|\beta_1| = 0$, resulting in an identity mapping $v = u$. Figure 1(b), (c), and (d) explicitly show that as $|\beta_1|$ approaches 1, $v = v(u)$ ($u \neq -\exp(\pi i/12)$) converges to a point $\beta_1/|\beta_1| = \exp(\pi i/12)$.

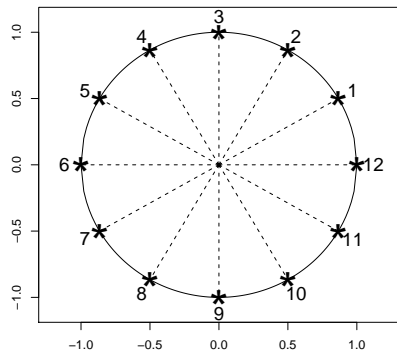
When $|\beta_1| = 1$, the mapping (2.1) maps the unit circle onto the point β_1 , i.e. $v = \beta_1$ for any u .

For the case of $|\beta_1| > 1$, (2.1) can be expressed as

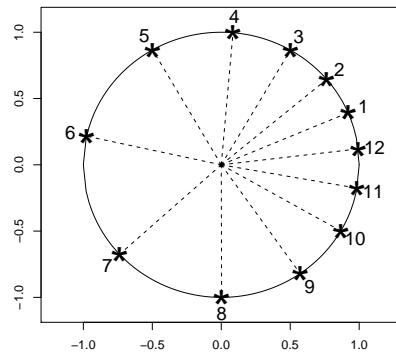
$$v = \beta_0 \frac{u + \beta_1}{1 + \overline{\beta_1}u} = \beta_0 \frac{u' + \beta_1'}{1 + \overline{\beta_1'}u'}, \quad (2.2)$$

where $u' = (\beta_1/|\beta_1|)(\beta_1\overline{u}/|\beta_1|)$ and $\beta_1' = 1/\overline{\beta_1}$. The above expression (2.2) shows that the Möbius transformation with $|\beta_1| > 1$ consists of two types of transformations, namely, reflection and the Möbius transformation with $|\beta_1'| < 1$, i.e.

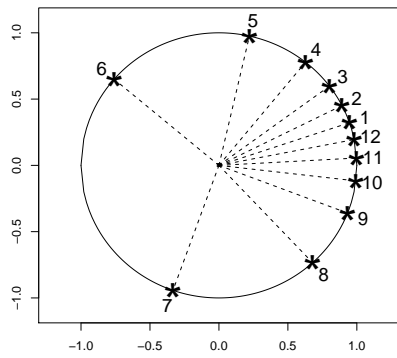
$$u \rightarrow (\beta_1/|\beta_1|)(\beta_1\overline{u}/|\beta_1|) \quad \text{and} \quad u \rightarrow \beta_0(u + \beta_1')/(1 + \overline{\beta_1'}u).$$



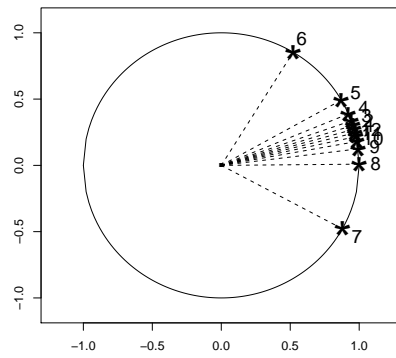
(a)



(b)



(c)



(d)

Figure 1. Plots of v for regression curve (2.1) for $u = \exp(2\pi ik/12)$, $k = 1, \dots, 12$, with $\beta_0 = 1$, $\arg(\beta_1) = \pi/12$, and: (a) $|\beta_1| = 0$; (b) $|\beta_1| = 0.3$; (c) $|\beta_1| = 0.6$; (d) $|\beta_1| = 0.9$.

2.2 Distribution for angular error

In this subsection we introduce a probability model for the angular error and give some known properties of the distribution.

Let z be a random variable on the unit circle in the complex plane. Then z has the wrapped Cauchy distribution or circular Cauchy distribution when the density for z is

$$f(z) = \frac{1}{2\pi} \frac{|1 - |\phi|^2|}{|z - \phi|^2}, \quad |z| = 1, \quad (2.3)$$

where $|\phi| \neq 1$. In this paper we extend the domain of the wrapped Cauchy distribution and define $z = \phi$ for $|\phi| = 1$. In the same way as McCullagh (1996), we write $z \sim C^*(\phi)$ to denote the wrapped Cauchy distribution (2.3).

By transforming z into polar co-ordinates $z = \exp(i\theta)$, $0 \leq \theta < 2\pi$, we obtain the density of θ , which is given by

$$f(\theta) = \frac{1}{2\pi} \frac{1 - \rho^2}{1 - 2\rho \cos(\theta - \mu) + \rho^2}, \quad 0 \leq \theta < 2\pi,$$

where

$$\mu = \arg(\phi) \quad \text{and} \quad \rho = \begin{cases} |\phi|, & |\phi| < 1, \\ 1/|\phi|, & |\phi| > 1. \end{cases}$$

It is clear that $\theta = \arg(\phi)$ for $|\phi| = 1$. Here μ is a mean direction and ρ a concentration of z or θ . The distribution is unimodal and symmetric about μ . When ρ is equal to 0, the distribution is the uniform distribution on the circle. As ρ tends to 1, the distribution approaches a point distribution with singularity at $z = \phi$ or $\theta = \mu$.

The properties of the wrapped Cauchy distribution are investigated, for example, by Mardia (1972) and McCullagh (1996). The following hold for the wrapped Cauchy distribution.

- (i) $z \sim C^*(\phi) \implies \beta_0 z \sim C^*(\beta_0 \phi), \quad |\beta_0| = 1,$
- (ii) $z_1 \sim C^*(\phi_1), z_2 \sim C^*(\phi_2), z_1 \perp z_2 \implies z_1 z_2 \sim C^*(\phi_1 \phi_2),$
- (iii) $z \sim C^*(\phi) \implies \frac{z + \beta_1}{1 + \overline{\beta_1} z} \sim C^*\left(\frac{\phi + \beta_1}{1 + \overline{\beta_1} \phi}\right), \quad \beta_1 \in \mathbb{C}.$

The properties (i) and (iii) show that if z is distributed as a uniform distribution $C^*(0)$, then the Möbius transformation of z generates the wrapped Cauchy distribution; i.e. $\beta_0(z + \beta_1)/(1 + \overline{\beta_1} z) \sim C^*(\beta_0 \beta_1)$ where $|\beta_0| = 1$ and $\beta_1 \in \mathbb{C}$.

Note that (ii) and (iii) do not hold for the von Mises distribution.

2.3 Definition and properties of the proposed regression model

This subsection provides a circular–circular regression model and investigates some properties of the model.

Let x be an independent variable which takes values on the unit circle in the complex plane and let y be the dependent variable. The complex parameters β_0 and β_1 are defined by $|\beta_0| = 1$ and $\beta_1 \in \mathbb{C}$. The proposed regression model is defined by

$$y = \beta_0 \frac{x + \beta_1}{1 + \overline{\beta_1} x} \varepsilon, \quad |x| = 1, \quad (2.4)$$

where $\varepsilon \sim C^*(\varphi)$, $0 \leq \varphi \leq 1$. Here we suppose $\arg(\varphi) = 0$ and $\varphi \leq 1$ since the mean direction of the angular error should be 0 and $C^*(\varphi) = C^*(1/\bar{\varphi})$ holds for any $\varphi \in \mathbb{C}$. We have already discussed the interpretation of β_0 and β_1 in §2.1. The parameter φ is the concentration or precision parameter. If $\varphi = 1$, then independent and dependent variables are correlated without error. The smaller the value of φ the more dependent the error variables. When $\varphi = 0$, the variable ε has a uniform distribution on the circle.

The conditional distribution of y given x is given by

$$y | x \sim C^* \left(\beta_0 \frac{x + \beta_1}{1 + \bar{\beta}_1 x} \varphi \right).$$

The following theorem holds for our regression model by applying well-known result in complex analysis. See Rudin (1987, Theorem 11.9) for the proof.

Theorem 1 *Suppose $f \equiv f(s, t)$ is a continuous real function on the closed unit disc satisfying $\nabla^2 f = \partial^2 f / \partial s^2 + \partial^2 f / \partial t^2 = 0$ for all $y = s + it$ in the open disc. Let g be a function which maps the unit circle in the complex plane into the closed disc in the complex plane. Then*

$$y | x \sim C^* \left(\beta_0 \frac{x + \beta_1}{1 + \bar{\beta}_1 x} \varphi \right) \implies E \{ f(y) | g(x) \} = f \left(\beta_0 \frac{g(x) + \beta_1}{1 + \bar{\beta}_1 g(x)} \varphi \right). \quad (2.5)$$

Using the result we obtain the mean direction and the concentration of $y | x$

$$\begin{aligned} \arg \{ E(y | x) \} &= \arg \left(\beta_0 \frac{x + \beta_1}{1 + \bar{\beta}_1 x} \right) = \arg(\beta_0 x) - 2 \arg(1 + \bar{\beta}_1 x), \\ |E(y | x)| &= \varphi. \end{aligned}$$

More generally, the k th trigonometric moment of $y | x$ is

$$E \left(y^k | x \right) = \left(\beta_0 \frac{x + \beta_1}{1 + \bar{\beta}_1 x} \varphi \right)^k. \quad (2.6)$$

Since the wrapped Cauchy distribution is closed under rotation and the Möbius transformation (see properties (i) and (iii) in Section 2.2), we obtain

$$\gamma_0 \frac{y + \gamma_1}{1 + \bar{\gamma}_1 y} \Big| x \sim C^* \left\{ \gamma_0 \frac{u(x) + \gamma_1}{1 + \bar{\gamma}_1 u(x)} \right\}, \quad (2.7)$$

where $|\gamma_0| = 1$, $\gamma_1 \in \mathbb{C}$, and $u(x) = \beta_0(x + \beta_1)\varphi / (1 + \bar{\beta}_1 x)$. Because of the fact that the linear fractional transformations form a group under composition, the parameter of the wrapped Cauchy (2.7) can also be expressed as the linear fractional transformation

$$\gamma_0 \frac{u(x) + \gamma_1}{1 + \bar{\gamma}_1 u(x)} = \frac{ax + b}{cx + d},$$

where $a = \gamma_0(\beta_0\varphi + \gamma_1\bar{\beta}_1)$, $b = \gamma_0(\gamma_1 + \beta_0\beta_1\varphi)$, $c = \bar{\beta}_1 + \bar{\gamma}_1\beta_0\varphi$, $d = 1 + \bar{\gamma}_1\beta_0\beta_1\varphi$. Our model is also closed for the bilinear fractional transformation of the independent variable

$$y \Big| \frac{\gamma_{00}x + \gamma_{01}}{\gamma_{10}x + \gamma_{11}} \sim C^* \left(\frac{ax + b}{cx + d} \right), \quad (2.8)$$

where $a = \beta_0(\gamma_{00} + \beta_1\gamma_{10}\varphi)$, $b = \beta_0(\gamma_{01} + \beta_1\gamma_{11})\varphi$, $c = \gamma_{10} + \overline{\beta_1}\gamma_{00}$, $d = \gamma_{11} + \overline{\beta_1}\gamma_{01}$, and

$$\begin{pmatrix} \gamma_{00} & \gamma_{01} \\ \gamma_{10} & \gamma_{11} \end{pmatrix} \in \text{GL}(2, \mathbb{C}).$$

If we assume that x is a random variable which has the wrapped Cauchy distribution $C^*(\phi)$, then the distribution of y is given by

$$y \sim C^* \left(\beta_0 \frac{\phi + \beta_1}{1 + \overline{\beta_1}\phi} \varphi \right). \quad (2.9)$$

The above is obvious from properties (i), (ii) and (iii) in Section 2.2.

2.4 Comparison with existing regression models

Let y be a dependent variable and z be a complex-valued nonstochastic covariate satisfying $\text{Im}(z) \neq 0$. The regression model of McCullagh (1996) is defined by

$$y | z \sim C \left(\frac{\beta_{00}z + \beta_{01}}{\beta_{10}z + \beta_{11}} \right),$$

where $C(\theta)$ is a Cauchy distribution on the real line with median $\text{Re}(\theta)$ and scale parameter $\text{Im}(\theta)$, and

$$\begin{pmatrix} \beta_{00} & \beta_{01} \\ \beta_{10} & \beta_{11} \end{pmatrix} \in \text{SL}(2, \mathbb{R}).$$

Although this model looks similar to ours at first glance, their model and ours are different. Their model is not circular–circular, but planar–linear regression model. In addition our model is obtained neither by wrapping $y | z$ nor by transforming $y' = (1 + iy)/(1 - iy)$, which are the transformations to generate a wrapped Cauchy distribution from a Cauchy distribution on the real line.

Our proposed regression model also has some relationship with the model of Fisher and Lee (1992) and Downs and Mardia (2002). Fisher and Lee (1992) proposed a linear–circular regression model in which the link function is expressed as a form of tangent function. Tangent function is also used in the link function of the circular–circular regression model of Downs and Mardia (2002). After some algebra, it is shown that our regression curve corresponds to their link function. However our model is different from theirs. The major distinction is the distribution for the angular error. In their model the probability model for the angular error has the von Mises distribution, whereas in our model we assume that the angular error is distributed as the wrapped Cauchy. Our model has some desirable properties that their model does not have. For example, the most important property, Theorem 1, does not hold for their model. In addition, all properties (2.6)–(2.9) do not hold for their model. These properties enable us to obtain the method of moments estimators and simpler expression of the Fisher information matrix and so forth. Furthermore, assuming the wrapped Cauchy as the angular error also makes it possible to determine properties of a related bivariate circular distribution and a multiple regression model for circular data, which will be discussed in Section 3 and Section 6.

3 Related Bivariate Circular Distribution

To our knowledge, no bivariate angular probability distribution has been used to model circular–circular regression. We now provide a bivariate circular distribution which could be helpful in modelling our circular–circular regression. It has the density

$$f(x, y) = \frac{1}{(2\pi)^2} \frac{|1 - \varphi^2|}{|y - \beta_0(x + \beta_1)\varphi/(1 + \overline{\beta_1}x)|^2} \frac{|1 - |\phi|^2|}{|x - \phi|^2}, \quad |x| = |y| = 1, \quad (3.1)$$

where $|\phi| \neq 1$, $0 \leq \varphi < 1$, and the other parameters are defined in the same way as in (2.4). The following properties hold for this distribution.

- (1) $y | x \sim C^* \left(\beta_0 \frac{x + \beta_1}{1 + \overline{\beta_1}x} \varphi \right)$,
- (2) $y \sim C^* \left(\beta_0 \frac{\phi + \beta_1}{1 + \overline{\beta_1}\phi} \varphi \right)$,
- (3) $x \sim C^*(\phi)$.

Hence, the marginal and conditional distributions are wrapped Cauchy distributions. The distribution (3.1) takes maximum value at $(x, y) = (\phi/|\phi|, \beta_0(x + \beta_1)\varphi/(1 + \overline{\beta_1}x))$ and minimum value at $(x, y) = (-\phi/|\phi|, -\beta_0(x + \beta_1)\varphi/(1 + \overline{\beta_1}x))$. For $|\beta_1| = 1$, x and y are independently distributed as $C^*(\phi)$ and $C^*(\beta_0\beta_1\varphi)$, respectively. The closer $|\beta_1|$ gets to 0, the more correlated x and y are. For $\varphi = 0$, x and y are independently distributed as $C^*(\phi)$ and the circular uniform distribution $C^*(0)$, respectively. The larger the value of φ , the greater the correlation between x and y . See Fisher and Lee (1983) for the definition of circular correlation.

4 Estimation and Test

4.1 Parameter estimation

Maximum likelihood estimation for the wrapped Cauchy distribution was investigated by Kent and Tyler (1988). However we cannot apply these results to the conditional distribution $y | x$ directly, since the mean direction is a function of the independent variable x . Therefore we need to obtain the maximum likelihood estimates of the wrapped Cauchy distribution in a different manner.

Let $y_j | x_j$ ($j = 1, \dots, n$) be a set of random samples from the wrapped Cauchy distribution $C^* \{ \beta_0(x_j + \beta_1)\varphi/(1 + \overline{\beta_1}x_j) \}$. The log-likelihood function for these samples is given by

$$\log L = C + \sum_{j=1}^n \{ \log |1 - \varphi^2| - 2 \log |y_j - \beta_0(x_j + \beta_1)\varphi/(1 + \overline{\beta_1}x_j)| \}.$$

Transform the independent and dependent variables by equating $(x_j, y_j) = (e^{i\theta_{1j}}, e^{i\theta_{2j}})$, and, for convenience, reparametrize $\beta_1 = re^{i\theta}$ ($r \geq 0$, $0 \leq \theta < 2\pi$) and $\beta'_0 = \arg(\beta_0)$.

Then the log-likelihood function can be expressed as

$$\log L = C + n \log(1 - \varphi^2) - \sum_{j=1}^n \log [1 - 2\varphi \cos \{\theta_{2j} - \mu(\theta_{1j})\} + \varphi^2], \quad (4.1)$$

where $\mu(\theta_{1j}) = \beta'_0 + \theta_{1j} - 2 \arg \{1 + re^{i(\theta_{1j} - \theta)}\}$.

When β_1 is known, the maximum likelihood estimates of β'_0 and φ are obtained by the recursive algorithm by Kent and Tyler (1988). The method of moments gives the estimators of β'_0 and φ as follows:

$$\hat{\beta}'_0 = \arg(C + iS) \quad \text{and} \quad \hat{\varphi} = \frac{1}{n} |C + iS|,$$

where $C = \sum_{j=1}^n \cos[\theta_{2j} - \theta_{1j} + 2 \arg\{1 + re^{i(\theta_{1j} - \theta)}\}]$ and $S = \sum_{j=1}^n \sin[\theta_{2j} - \theta_{1j} + 2 \arg\{1 + re^{i(\theta_{1j} - \theta)}\}]$.

4.2 Fisher information matrix

Using the log-likelihood for $(\beta'_0, r, \theta, \varphi)$ given by (4.1). We find that

$$-E \left(\frac{\partial^2}{\partial r \partial \varphi} \log L \right) = -E \left(\frac{\partial^2}{\partial \theta \partial \varphi} \log L \right) = -E \left(\frac{\partial^2}{\partial \beta'_0 \partial \varphi} \log L \right) = 0.$$

Hence, φ and (r, θ, β'_0) are orthogonal. The other elements of the Fisher information matrix are calculated as

$$\begin{aligned} E \left\{ \left(\frac{\partial}{\partial \varphi} \log L \right)^2 \right\} &= \frac{2n}{(1 - \varphi^2)^2}, \\ E \left\{ \left(\frac{\partial}{\partial r} \log L \right) \left(\frac{\partial}{\partial \beta'_0} \log L \right) \right\} &= \frac{2\varphi^2}{(1 - \varphi^2)^2} \sum_{j=1}^n \frac{\partial \mu_j}{\partial r}, \\ E \left\{ \left(\frac{\partial}{\partial r} \log L \right) \left(\frac{\partial}{\partial \theta} \log L \right) \right\} &= \frac{2\varphi^2}{(1 - \varphi^2)^2} \sum_{j=1}^n \frac{\partial \mu_j}{\partial r} \frac{\partial \mu_j}{\partial \theta}, \\ E \left\{ \left(\frac{\partial}{\partial \beta'_0} \log L \right) \left(\frac{\partial}{\partial \theta} \log L \right) \right\} &= \frac{2\varphi^2}{(1 - \varphi^2)^2} \sum_{j=1}^n \frac{\partial \mu_j}{\partial \theta}, \\ E \left\{ \left(\frac{\partial}{\partial r} \log L \right)^2 \right\} &= \frac{2\varphi^2}{(1 - \varphi^2)^2} \sum_{j=1}^n \left(\frac{\partial \mu_j}{\partial r} \right)^2, \\ E \left\{ \left(\frac{\partial}{\partial \theta} \log L \right)^2 \right\} &= \frac{2\varphi^2}{(1 - \varphi^2)^2} \sum_{j=1}^n \left(\frac{\partial \mu_j}{\partial \theta} \right)^2, \\ E \left\{ \left(\frac{\partial}{\partial \beta'_0} \log L \right)^2 \right\} &= \frac{2n\varphi^2}{(1 - \varphi^2)^2}, \end{aligned}$$

where

$$\frac{\partial \mu_j}{\partial \theta} = \frac{2r \{r + \cos(\theta_{1j} - \theta)\}}{1 + 2r \cos(\theta_{1j} - \theta) + r^2}, \quad \frac{\partial \mu_j}{\partial r} = \frac{-2 \sin(\theta_{1j} - \theta)}{1 + 2r \cos(\theta_{1j} - \theta) + r^2}.$$

4.3 A test of independence

To investigate if the model (2.4) provides a better fit than the independence model, we test $H_0 : r = 1$ against $H_1 : r \neq 1$. The likelihood ratio test gives the test statistic as

$$T = -2 \log \frac{\max L_0}{\max L_1},$$

where $\max L_0 = \max_{\varphi, \beta'_0} L_0(\varphi, \beta'_0, \theta = 0)$ and $\max L_1 = \max_{\varphi, r, \theta, \beta'_0} L_1(\varphi, r, \theta, \beta'_0)$. Under the null hypothesis, T is asymptotically distributed as a chi-square distribution with one degree of freedom. Here $\max L_0$ is easily obtained using the algorithm of Kent and Tyler (1988). We reject the null hypothesis when T is large.

Other large sample theories, such as Wald test and score test, could also be useful for inference for the proposed model.

5 Examples

Example 1. In a marine biology study by Dr. Robert R. Warner at University of California, Santa Barbara (Lund, 1999), whether the spawning time of a particular fish (T_S) depends on the time of the low tide (T_{LT}) is of interest. The data were gathered in St. Croix, the U.S. Virgin Islands. To study the dependence of T_S on T_{LT} , we converted the period 0 to 20 hours of T_{LT} to $[0, 2\pi)$. Raw data of T_S range from 11.76 to 14.98 hours, and we converted the period 11.5 to 15 hours of T_S to $\frac{11.5}{24} \times 2\pi + [0, 2\pi) \pmod{2\pi}$. Paired T_S and T_{LT} are thus bivariate circular data, and they are plotted as circles in Figure 2(a). In the following, we apply model (2.4) to investigate whether and how T_S depends on T_{LT} .

Table 1. Maximum likelihood estimates of the parameters, the maximum log-likelihood, and AIC of model (2.4).

Model	β'_0	r	θ	ϕ	Log-likelihood	AIC
(2.4)	1.77	0.552	0.506	0.397	-169.2	346.4

The maximum likelihood estimates, the maximum log-likelihood, and AIC of each model are given in Table 1. The test of independence for model (2.4) yields the test statistic $T = -2\{(-182.0) - (-169.2)\} = 25.6$. This test is highly significant and the assumption of independence is rejected.

Figure 2(b) exhibits Q-Q plot for model (2.4), which is a plot of quantiles of wrapped Cauchy (x -axis) and quantiles of empirical distributions (y -axis). This plot seems to show that residual is distributed as wrapped Cauchy distribution. Figure 2(c) displays the plot of circular distance. Here the circular distance is defined by $d(y, \hat{y}) = 1 - \cos(y - \hat{y})$ where y is a dependent variable and \hat{y} is a predictor in radians given by $\hat{y} = \hat{\beta}'_0 + x - 2 \arg\{1 + re^{i(x-\hat{\theta})}\}$. Observation numbers of outliers are marked on the plot. The figure shows that the circular distance of most observations lie in $[0, 1]$.

Example 2. For another illustrative example, we consider a dataset of wind directions. The wind direction at 6 a.m. and 12 noon was measured each day at a weather station in

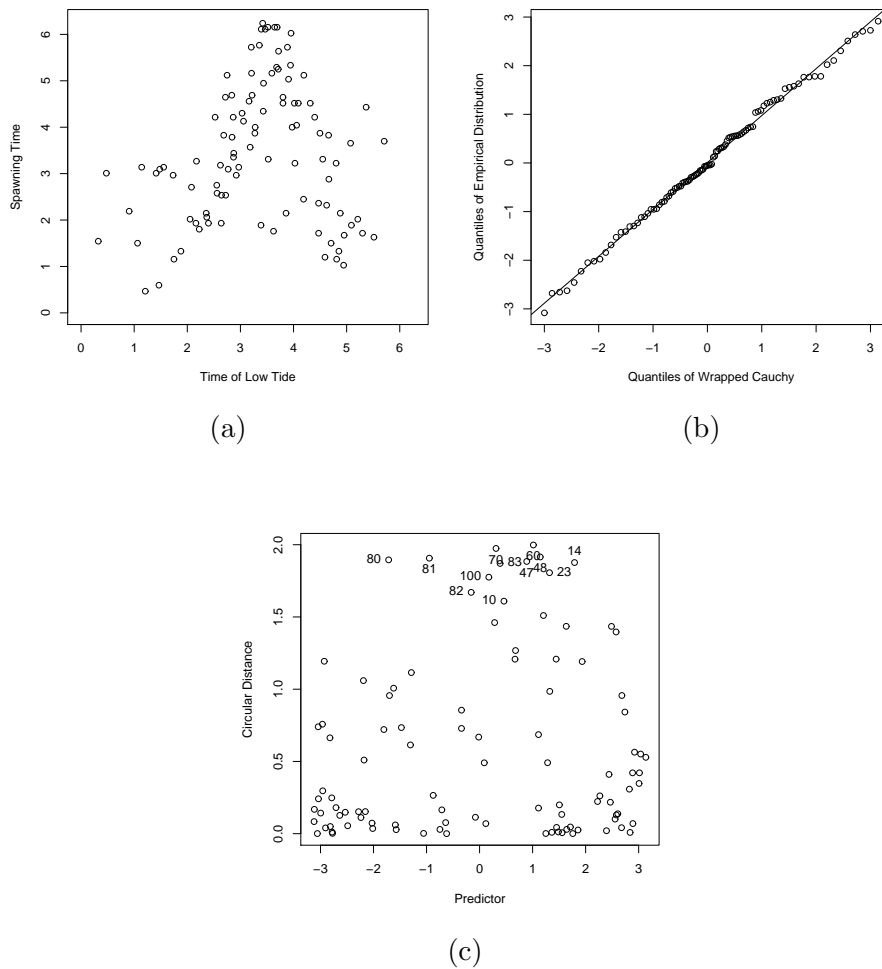


Figure 2. (a) planar plot of the spawning time of certain fish and the low tide time. Both time are converted into angles $[0, 2\pi)$. (b) Q-Q plot for model (2.4). Quantiles of wrapped Cauchy (x -axis) and quantiles of empirical distributions (y -axis) are plotted. For visual clearness, angles are transformed from $[0, 2\pi)$ into $[-\pi, \pi)$. (c) Plots of circular distance for model (2.4). Observation numbers are marked for outliers.

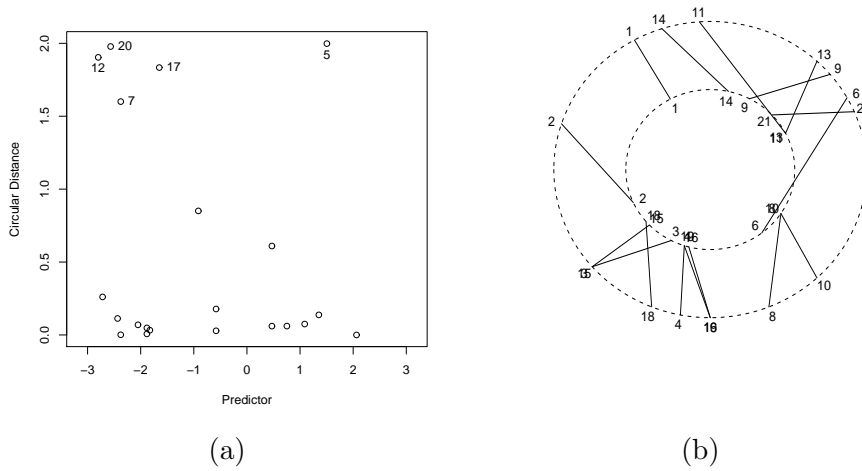


Figure 3. (a) plot of circular distance, and (b) plot of predictors and dependent variables, in which the predictors are plotted on the smaller circle whereas the dependent variables are marked on the larger one.

Milwaukee for 21 consecutive days. The data are taken from Table B.21 of Fisher (1993). We use model (2.4) for regressing the wind direction at 12 noon on that at 6 a.m.

Table 2. Maximum likelihood estimates of the parameters, the maximum log-likelihood, and AIC of model (2.4).

Model	β'_0	r	θ	ϕ	Log-likelihood	AIC
(2.4)	1.27	0.528	2.59	0.550	-32.3	72.6

Table 2 shows the maximum likelihood estimates of the parameters, maximum log-likelihood, and AIC of the model. Judging from the AIC, model (2.4) provides a better fit than the Downs and Mardia model, whose AIC is 74.4. The test of independence for the model (2.4) in Section 4.3 yields the test statistic as $T = -2\{(-38.5) - (-32.3)\} = 12.4$. This test is highly significant and the assumption of independence is rejected.

The plot of the circular distance is given in Figure 3(a). The observation numbers of outliers are marked on the plot. This plot shows that the outliers are observed in 5, 7, 12, 17 and 20. Apart from those five outliers, model (2.4) seems to provide a satisfactory fit to the data.

Finally, the predictors and dependent variables except for outliers are plotted by observation numbers in Figure 3(b). The plots on the larger circle refer to the dependent variables, while those on the smaller one are the predictors from model (2.4). The short line between the predictor and dependent variable means good fit of the model to the observation. Judging from Figure 3(b), our model seems to provide satisfactory fit to the data. For the interpretation of how the dependent variables are transformed through the Möbius transformation, see Section 2.1.

6 Circular Multiple Regression

Downs and Mardia (2002) presented a recursive method for regressing the dependent angle on multiple independent angles. Here we propose another method for circular multiple regression by extending the model (2.4). Let x_1, \dots, x_p be independent variables on the unit circle in the complex plane. Suppose that y is a dependent variable. The proposed circular multiple regression is defined by

$$y = \beta_0 \prod_{j=1}^p \frac{x_j + \beta_j}{1 + \overline{\beta_j} x_j} \varepsilon, \quad (6.1)$$

where $|\beta_0| = 1$, $\beta_j \in \mathbb{C}$, $j = 1, \dots, p$ are parameters. The angular error is distributed as the wrapped Cauchy $\varepsilon \sim C^*(\varphi)$, where $0 \leq \varphi < 1$. The parameter φ works as a concentration or precision parameter and β_0 as a rotation parameter. The other parameters β_1, \dots, β_p have the same interpretation as β_1 in the bivariate model. When $p = 1$, the model (6.1) coincides with (2.4). The conditional distribution of y given $x = (x_1, \dots, x_p)'$ is

$$y | x \sim C^* \left(\beta_0 \prod_{j=1}^p \frac{x_j + \beta_j}{1 + \overline{\beta_j} x_j} \varphi \right).$$

Many of the properties for the bivariate model (2.4) also hold for the model (6.1). For instance, we can get the trigonometric moments of $y | x$ by (2.5). The mean direction and the concentration of $y | x$ also have simple forms. The Fisher information matrix is calculated similarly to that of model (2.4). We can easily show that φ and the other parameters β_0, \dots, β_p are orthogonal. Estimation and inference for the model are given in much the same way as before using numerical methods and large sample theories, respectively.

When we assume that the x_i are random samples from $C^*(\phi_i)$, $i = 1, \dots, p$, then y is distributed as $C^*[\beta_0 \prod_{j=1}^p \{(\phi_j + \beta_j)/(1 + \overline{\beta_j} \phi_j)\} \varphi]$. This follows from properties (i), (ii) and (iii) in Section 2.2.

7 Discussion

Circular–circular regression is useful for analyzing bivariate circular data. Among existing regression models, the *raison d'être* of our model is its tractability and expandability. The tractability derives from the theory of the Möbius transformation and the wrapped Cauchy distribution. As discussed in Section 2.2 in this paper, the wrapped Cauchy is related to the Möbius transformation, and thus enables us to obtain a number of desirable properties for our model. As for the expandability, our regression model could provide some topics to other related fields. For example, the related bivariate circular distribution and multiple circular regression model, which are briefly discussed in Section 3 and Section 6, could be possible fields worth carrying out further research.

Acknowledgment

The authors are grateful to Dr. Nicholas I. Fisher and Dr. Peter J. Thomson for helpful comments on our paper. This work was supported in part by the Ministry of Education, Culture, Sport, Science and Technology in Japan under a Grant-in-Aid of the 21st Century Center of Excellence for Integrative Mathematical Sciences: Progress in Mathematics Motivated by Social and Natural Sciences.

References

- Downs, T. D. (2003). Spherical regression. *Biometrika* **90**, 655-668.
- Downs, T. D. and Mardia, K. V. (2002). Circular regression. *Biometrika* **89**, 683-697.
- Fisher, N. I. (1993). *Statistical Analysis of Circular Data*. Cambridge University Press, Cambridge.
- Fisher, N. I. and Lee, A. J. (1983). A correlation coefficient for circular data. *Biometrika* **70**, 327-332.
- Fisher, N. I. and Lee, A. J. (1992). Regression models for an angular response. *Biometrics* **48**, 665-677.
- Jones, M. C. (2004). The Möbius distribution on the disc. *Ann. Inst. Statist. Math.* **56**, 733-742.
- Kent, J. T. and Tyler, D. E. (1988). Maximum likelihood estimation for the wrapped Cauchy distribution. *J. Appl. Statist.* **15**, 247-254.
- Kolassa, J. and McCullagh, P. (1990). Edgeworth series for lattice distributions. *Ann. Statist.* **18**, 981-985.
- Lund, U. J. (1999). Least circular distance regression for directional data. *J. Appl. Statist.* **26**, 723-733.
- Mardia, K. V. (1972). *Statistics of Directional Data*. Academic Press, New York.
- Mardia, K. V. and Jupp, P. E. (2000). *Directional Statistics*. Wiley, Chichester.
- McCullagh, P. (1989). Some statistical properties of a family of continuous univariate distributions. *J. Am. Statist. Assoc.* **84**, 125-129.
- McCullagh, P. (1996). Möbius transformation and Cauchy parameter estimation. *Ann. Statist.* **24**, 787-808.
- Minh, D. L. P. and Farnum, N. R. (2003). Using bilinear transformations to induce probability distributions. *Commun. Statist. B* **32**, 1-9.
- Rivest, L.-P. (1997). A decentred predictor for circular-circular regression. *Biometrika* **84**, 717-726.

Rudin, W. (1987). *Real and Complex Analysis*. McGraw-Hill, New York.

Seshadri, V. (1991). A family of distributions related to the McCullagh family. *Statist. Prob. Lett.* **12**, 373-378.

Department of Mathematics
Faculty of Science and Technology
Keio University

Research Report

2005

- [05/001] T. Kawazoe, *On the Littlewood-Paley g function and the Lusin area function on real rank 1 semisimple Lie groups*, KSTS/RR-05/001, March 28, 2005
- [05/002] A. Sako, T. Suzuki, *Partition functions of supersymmetric gauge theories in noncommutative \mathbb{R}^{2D} and their unified perspective*, KSTS/RR-05/002, April 25, 2005
- [05/003] K. Aoyama, K. Shimizu, *A generalization of the inverse trinomial*, KSTS/RR-05/003, June 23, 2005
- [05/004] T. Tanaka, *Algebraic independence of modified reciprocal sums of products of Fibonacci numbers*, KSTS/RR-05/004, July 27, 2005
- [05/005] Y. Aoki, T. Kato, *New method of ground reflection coefficient reconstruction and shape recovery by another SAR signal modeling*, KSTS/RR-05/005, October 12, 2005
- [05/006] A. Sako, T. Suzuki, *Dimensional reduction of Seiberg-Witten monopole equations, $N=2$ noncommutative supersymmetric field theories and Young diagrams*, KSTS/RR-05/006, November 28, 2005

2006

- [06/001] N. Kumasaka, R. Shibata, *High dimensional data visualisation: Textile plot*, KSTS/RR-06/001, February 13, 2006
- [06/002] H. Omori, Y. Maeda, N. Miyazaki, A. Yoshioka, *Geometric objects in an approach to quantum geometry*, KSTS/RR-06/002, March 15, 2006
- [06/003] M. Nakasuji, *Prime geodesic theorem for higher dimensional hyperbolic manifold*, KSTS/RR-06/003, March 15, 2006
- [06/004] T. Kawazoe, *Uncertainty principle for the Fourier-Jacobi transform*, KSTS/RR-06/004, April 11, 2006
- [06/005] K. Kikuchi, S. Ishikawa, *Regression analysis, Kalman filter and measurement error model in measurement theory*, KSTS/RR-06/005, April 19, 2006
- [06/006] S. Kato, K. Shimizu, G. S. Shieh, *A circular-circular regression model*, KSTS/RR-06/006, May 24, 2006