Statistica Sinica **18**(2008), 633-645

A CIRCULAR–CIRCULAR REGRESSION MODEL

Shogo Kato, Kunio Shimizu and Grace S. Shieh

Institute of Statistical Mathematics, Keio University and Academia Sinica

Abstract: This paper provides a regression model in which both covariates and responses are angular variables. The regression curve is expressed as a form of the Möbius circle 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, including estimation and testing procedures, are obtained. The proposed methods are applied to marine biology and wind direction data.

Key words and phrases: Bivariate circular distribution, Möbius transformation, wrapped Cauchy distribution.

1. Introduction

Some regression models in which both covariates and responses 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 complex plane onto itself. With some restrictions on the parameters, this mapping maps, for example, the unit circle onto itself or the unit circle onto the real line. One of the earlier works in directional statistics in which the Möbius transformation appeared was McCullagh (1996). In this paper he discussed the connection between the real 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 disc which is generated by applying the Möbius transformation to the symmetric beta or Pearson type II distribution. McCullagh (1989) and Seshadri (1991) transformed their distributions via a one-to-one mapping which has the same form as the Möbius transformation and maps the interval $(-1, 1)$ onto itself.

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 circle transformation, and 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 direction data.

2. Circular Regression Model

Let responses $y_1, \ldots, y_n$ be independent, and let $x_1, \ldots, x_n$ be nonstochastic covariates which take values on the unit circle, $\Omega = \{z \in \mathbb{C}; |z| = 1\}$, in the complex plane. In the proposed regression model, the conditional distribution of y_j given $x_1, \ldots, x_n$ has the wrapped Cauchy distribution with mean direction $\arg\{v(x_j)\}$ defined in Section 2.1 and concentration $\varphi \in [0, 1]$.

In Section 2.1 we define the regression curve v and investigate its properties. The wrapped Cauchy distribution for the regression error and some properties of the regression model are discussed in Section 2.2 and Section 2.3, respectively. Comparison with existing regression models is given in Section 2.4.

2.1. Regression curve

Suppose β_0 and β_1 are complex parameters with $\beta_0 \in \Omega$ and $\beta_1 \in \mathbb{C}$. The regression curve of the proposed regression model is defined by

$$v = v(x; \beta_0, \beta_1) = \beta_0 \frac{x + \beta_1}{1 + \overline{\beta_1} x}, \quad x \in \Omega, \tag{2.1}$$

where the mapping with $|\beta_1| \neq 1$ is called a Möbius circle transformation, it is a one-to-one mapping which carries the unit circle onto itself.

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

(1) Translations: $z \rightarrow z + b$,
(2) Rotations: $z \rightarrow az$, $a \in \Omega$,
(3) Homotheties: $z \rightarrow rz$, $r > 0$,
(4) Inversion: $z \rightarrow 1/z$.

Note that these transformations exhibit the action of the group on the complex plane, not on the circle. For $\beta_1 \neq 0$, v can be expressed as

$$v = \beta_0\Big(\frac{1}{\overline{\beta_1}} + \frac{\lambda}{\overline{\beta_1}x + 1}\Big), \quad \lambda = \beta_1 - \frac{1}{\overline{\beta_1}}.$$

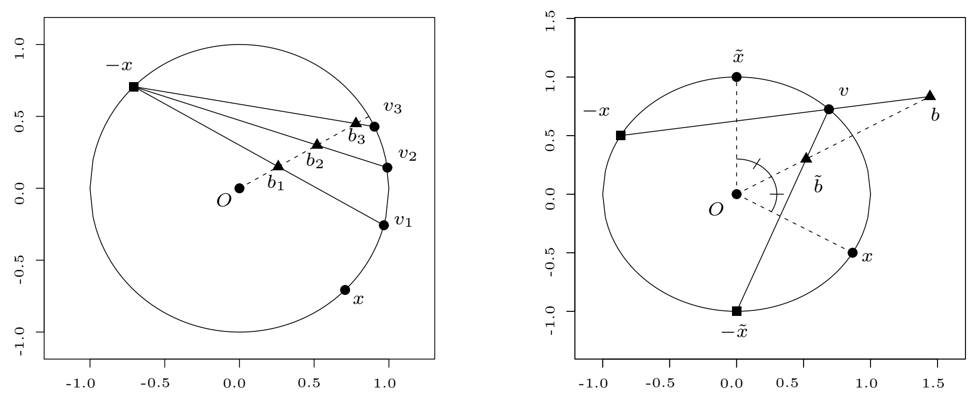
In (2.1), β_0 is evidently a rotation parameter, but the interpretation of β_1 is more complicated. However, the function of β_1 in (2.1) for $|\beta_1| < 1$ is revealed as follows. Assume, without loss of generality, that $\beta_0 = 1$. Then, for any $\beta_1 \in \mathbb{C}$ and any $x \in \Omega$, (2.1) implies that β_1 is the projection point for the straight line projection of $-x$ on the unit circle to the point v on the unit circle. From this fact, 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 $x = -\beta_1/|\beta_1|$, which is invariant under the Möbius circle transformation for any $|\beta_1| < 1$.

Figure 1(a) exhibits the behaviour of (2.1) for some specified values of β_1 for $|\beta_1| < 1$. Figure 1(a) explicitly shows that as $|\beta_1|$ approaches 1, $v(x \neq -\exp(\pi i/12))$ converges to a point $\beta_1/|\beta_1| = \exp(\pi i/12)$. It is also clear from the figure that as $|\beta_1|$ tends to 0, v approaches the identity mapping. When $|\beta_1| = 1$, the mapping (2.1) maps the unit circle onto the point β_1, i.e., $v = \beta_1$ for any x. For the case of $|\beta_1| > 1$, (2.1) can be expressed as

$$v = \beta_0\frac{x + \beta_1}{1 + \overline{\beta_1}x} = \beta_0\frac{\tilde{x} + \tilde{\beta}_1}{1 + \overline{\tilde{\beta}_1}\tilde{x}}, \tag{2.2}$$

where $\tilde{x} = (\beta_1/|\beta_1|)(\beta_1\overline{x}/|\beta_1|)$ and $\tilde{\beta}_1 = 1/\overline{\beta_1}$. The expression (2.2) shows that the Möbius circle transformation with $|\beta_1| > 1$ consists of two types of transformations, namely, reflection and the Möbius circle transformation with $|\tilde{\beta}_1| < 1$, i.e.,

$$x \longmapsto \Big(\frac{\beta_1}{|\beta_1|}\Big)\Big(\frac{\beta_1\overline{x}}{|\beta_1|}\Big) \quad \text{and} \quad x \longmapsto \beta_0\frac{x + \tilde{\beta}_1}{1 + \overline{\tilde{\beta}_1}x}.$$



(a) (b)

Figure 1. (a) plot of $v(x;\beta_0,\beta_1)$ for regression curve (2.1) for $x=\exp(-\pi i/4)$ with $\beta_0=1$, $\arg(\beta_1)=\pi/6$ and: $|\beta_1|=0.3$; $|\beta_1|=0.6$; $|\beta_1|=0.9$. Points on the plot are defined by $v_j=v(x;1,b_j)$, $b_j=0.3j\exp(\pi i/6)$, $j=1,2,3$. (b) plot of v, x, $\tilde{x}$, β_1, $\tilde{\beta}_1$ for equation (2.2) for $\beta_0=1$, $x=\exp(-\pi i/6)$, $\beta_1=5\exp(\pi i/6)/3$. Parameters β_1 and $\tilde{\beta}_1$ are expressed as b and $\tilde{b}$ on the plot, respectively.

Figure 1(b) displays an example in which (2.2) holds for selected values of β_0,β_1 and x. The figure clearly shows the fact that the Möbius circle transformation, with $|\beta_1|>1$, is made up of the two transformations mentioned above.

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

$$f(y)=\frac{1}{2\pi}\frac{|1-|\phi|^2|}{|y-\phi|^2},\quad y\in\Omega, \tag{2.3}$$

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

By transforming y and ϕ into polar co-ordinates $y=\exp(i\theta)$, $\phi=\rho\exp(i\mu)$, $0\le\theta$, $\mu<2\pi$, we obtain the density of θ, which is given by

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

where

$$\rho = \begin{cases} |\phi|, & |\phi| < 1, \\ \dfrac{1}{|\phi|}, & |\phi| > 1. \end{cases}$$

It is clear that $\mu = \arg(\phi)$ for $\phi \in \Omega$. Here μ is a mean direction and ρ a concentration of y 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 $y = \phi$ or $\theta = \mu$.

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

(i) $y \sim C^*(\phi) \Longrightarrow \beta_0 y \sim C^*(\beta_0\phi)$, $\beta_0 \in \Omega$,

(ii) $y_1 \sim C^*(\phi_1)$, $y_2 \sim C^*(\phi_2)$, $y_1 \perp y_2$, $|\phi_1|, |\phi_2| \le 1 \Longrightarrow y_1 y_2 \sim C^*(\phi_1\phi_2)$,

(iii) $y \sim C^*(\phi) \Longrightarrow \dfrac{y+\beta_1}{1+\overline{\beta_1}y} \sim C^*\Big(\dfrac{\phi+\beta_1}{1+\overline{\beta_1}\phi}\Big)$, $\beta_1 \in \mathbb{C}$,

(iv) $y \sim C^*(\phi) \Longrightarrow y \sim C^*(1/\overline{\phi})$.

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

Note that (ii)$-$(iv) do not hold for the von Mises distribution.

2.3. Some properties of the proposed regression model

This subsection discusses some properties of the proposed regression model. For simplicity of expression, we consider a case in which a single pair of a covariate and a response is observed.

Let x be a covariate which takes values on the unit circle in the complex plane, and let y be a response. The complex parameters are $\beta_0 \in \Omega$ and $\beta_1 \in \mathbb{C}$. The proposed regression model is given by

$$y = \beta_0 \frac{x+\beta_1}{1+\overline{\beta_1}x}\varepsilon, \quad x \in \Omega, \tag{2.4}$$

where $\varepsilon \sim C^*(\varphi)$, $0 \le \varphi \le 1$. Here the restriction on the domain of φ is valid because the mean direction of the angular error should be 0, and $C^*(\varphi) = C^*(1/\overline{\varphi})$ holds for any $\varphi \in \mathbb{C}$. We have already discussed the interpretation of β_0 and β_1 in Section 2.1. The parameter φ is the concentration or precision parameter. If $\varphi = 1$, then covariates and responses are correlated without error. The smaller the value of φ, the less concentrated the error variables. When $\varphi = 0$, the variable ε has a uniform distribution on the circle.

The conditional distribution of y given x is

$$y|x \sim C^*(\phi_{y|x}) \text{ where } \phi_{y|x} = \exp(i\mu_{y.x})\varphi \text{ and } \mu_{y.x} = \arg\Big(\beta_0 \frac{x+\beta_1}{1+\overline{\beta_1}x}\Big). \quad (2.5)$$

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

Theorem 1. *If $y \sim C^*(\phi)$ where $|\phi| \le 1$, then $E\{g(y)\} = g(\phi)$ for any mapping g on the closed unit disc which is continuous on the closed unit disc and analytic on the open unit disc.*

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

$$\arg\{E(y|x)\} = \mu_{y.x} = \arg(\beta_0 x) - 2\arg(1+\overline{\beta_1}x), \quad |E(y|x)| = \varphi.$$

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

$$E(y^k|x) = \phi_{y|x}^k. \quad (2.6)$$

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

$$\gamma_0 \frac{y+\gamma_1}{1+\overline{\gamma_1}y} \,\Big|\, x \sim C^*\Big(\gamma_0 \frac{\phi_{y|x}+\gamma_1}{1+\overline{\gamma_1}\phi_{y|x}}\Big), \quad (2.7)$$

where $\gamma_0 \in \Omega$, $\gamma_1 \in \mathbb{C}$. 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

$$\gamma_0 \frac{\phi_{y|x}+\gamma_1}{1+\overline{\gamma_1}\phi_{y|x}} = \frac{a_{00}x + a_{01}}{a_{10}x + a_{11}},$$

where $a_{00} = \gamma_0(\beta_0\varphi + \gamma_1\overline{\beta_1})$, $a_{01} = \gamma_0(\gamma_1 + \beta_0\beta_1\varphi)$, $a_{10} = \overline{\beta_1} + \overline{\gamma_1}\beta_0\varphi$, $a_{11} = 1 + \overline{\gamma_1}\beta_0\beta_1\varphi$.

Although property (2.7) is mathematically attractive, it is remarked here that the absolute values of the parameters in (2.7) depend on x and therefore homoscedasticity no longer holds. This formulation should be avoided unless heteroscedasticity is desired. To avoid this heteroscedasticity, one can transform y to $w = \gamma_0\{(y+\gamma_1)/(1+\overline{\gamma_1}y)\}$ and then use the model set up by (2.4) and (2.5) for $w|x$. Similarly, the following property holds for the Möbius circle transformation of the covariate:

$$y \,\Big|\, \gamma_0 \frac{x+\gamma_1}{1+\overline{\gamma_1}x} \sim C^*\Big(\frac{b_{00}x + b_{01}}{b_{10}x + b_{11}}\Big), \quad (2.8)$$

where $\gamma_0 \in \Omega$, $\gamma_1 \in \mathbb{C}$, $b_{00} = \beta_0(1 + \overline{\gamma_1}\beta_1)\varphi$, $b_{01} = \beta_0(\gamma_1 + \beta_1)\varphi$, $b_{10} = \overline{\gamma_1} + \overline{\beta_1}$, $b_{11} = 1 + \gamma_1\overline{\beta_1}$.

If we assume that x is a random variable which has the wrapped Cauchy distribution $C^*(\phi)$ and is independent of ε in (2.4), then the distribution of y is given by

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

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

2.4. Comparison with existing regression models

McCullagh (1996, Equation 28) proposed a regression model in which the error is assumed to follow a Cauchy distribution on the real line. Although his model looks similar to ours at first glance, his model and ours are different. His 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 models 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. The tangent function is also used as 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 angular error assumes 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, such as Theorem 1 and properties (2.6)−(2.9).

3. Related Bivariate Circular Distribution

To our knowledge, no bivariate angular 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

$$f(x, y) = \frac{1}{(2\pi)^2} \frac{|1 - \varphi^2|}{|y - \phi_{y|x}|^2} \frac{|1 - |\delta|^2|}{|x - \delta|^2}, \quad x, y \in \Omega, \tag{3.1}$$

where $|\delta| \neq 1$ and the other parameters are defined as in (2.4) and (2.5). The following properties hold for this distribution:

(1) $y|x \sim C^*(\phi_{y|x})$,

(2) $y \sim C^*\Big(\beta_0 \frac{\delta + \beta_1}{1 + \overline{\beta_1}\delta}\varphi\Big)$,

(3) $x \sim C^*(\delta)$.

Hence, the marginals and the conditional of y given x are wrapped Cauchy distributions. The distribution (3.1) takes maximum (minimum) value for each x at $y = \exp(i\mu_{y.x})(\exp(-i\mu_{y.x}))$. For $\beta_1 \in \Omega$, x and y are independently distributed as $C^*(\delta)$ and $C^*(\beta_0\beta_1\varphi)$, respectively. The closer $|\beta_1|$ gets to 0, the closer $\exp(i\mu_{y.x})$ is to being a pure rotation of x. For $\varphi = 0$, x and y are independently distributed as $C^*(\delta)$ 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 a 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 covariate x. Therefore we need to obtain the maximum likelihood estimates of the wrapped Cauchy distribution in a different manner.

Let $y_j|x_j$ $(j = 1, \ldots, n)$ be a set of random samples from the wrapped Cauchy distribution $C^*(\phi_{y_j|x_j})$. The log-likelihood function for these samples is

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

Transform the covariates and responses by taking $(x_j, y_j) = (e^{i\theta_{x_j}}, e^{i\theta_{y_j}})$ and, for convenience, reparametrize (β_0, β_1) as $(e^{i\theta_0}, re^{i\theta_1})$, where $r > 0$, $0 \le \theta_0$, $\theta_1 < 2\pi$. Then the log-likelihood function can be expressed as

$$\log L = C + n\log(1-\varphi^2) - \sum_{j=1}^{n} \log\left\{1 - 2\varphi\cos(\theta_{y_j} - \mu_{y_j|x_j}) + \varphi^2\right\}, \qquad (4.1)$$

where $\mu_{y_j|x_j} = \theta_0 + \theta_{x_j} - 2\arg\{1 + re^{i(\theta_{x_j}-\theta_1)}\}$.

If β_1 is known, the maximum likelihood estimates of θ_0 and φ are obtained by the recursive algorithm by Kent and Tyler (1988). The method of moments gives the estimators of θ_0 and φ as follows:

$$\hat{\theta}_0 = \arg(C_n + iS_n) \quad \text{and} \quad \hat{\varphi} = \frac{1}{n}|C_n + iS_n|,$$

where $C_n = \sum_{j=1}^{n} \cos[\theta_{y_j} - \theta_{x_j} + 2\arg\{1 + re^{i(\theta_{x_j}-\theta_1)}\}]$ and $S_n = \sum_{j=1}^{n} \sin[\theta_{y_j} - \theta_{x_j} + 2\arg\{1 + re^{i(\theta_{x_j}-\theta_1)}\}]$.

4.2. Fisher information matrix

Using the log-likelihood for $(\theta_0, r, \theta_1, \varphi)$ given by (4.1), we find that

$$-E\Big(\frac{\partial^2}{\partial\theta_0\partial\varphi}\log L\Big) = -E\Big(\frac{\partial^2}{\partial r\partial\varphi}\log L\Big) = -E\Big(\frac{\partial^2}{\partial\theta_1\partial\varphi}\log L\Big) = 0.$$

Hence, φ and (θ_0, r, θ_1) are orthogonal. The other elements of the Fisher information matrix are calculated as

$$E\Big\{\Big(\frac{\partial}{\partial\theta_0}\log L\Big)^2\Big\} = \frac{2n\varphi^2}{(1-\varphi^2)^2},$$

$$E\Big\{\Big(\frac{\partial}{\partial r}\log L\Big)^2\Big\} = \frac{2\varphi^2}{(1-\varphi^2)^2}\sum_{j=1}^{n}\Big(\frac{\partial\mu_{y_j|x_j}}{\partial r}\Big)^2,$$

$$E\Big\{\Big(\frac{\partial}{\partial\theta_1}\log L\Big)^2\Big\} = \frac{2\varphi^2}{(1-\varphi^2)^2}\sum_{j=1}^{n}\Big(\frac{\partial\mu_{y_j|x_j}}{\partial\theta_1}\Big)^2,$$

$$E\Big\{\Big(\frac{\partial}{\partial\varphi}\log L\Big)^2\Big\} = \frac{2n}{(1-\varphi^2)^2},$$

$$E\Big\{\Big(\frac{\partial}{\partial\theta_0}\log L\Big)\Big(\frac{\partial}{\partial r}\log L\Big)\Big\} = \frac{2\varphi^2}{(1-\varphi^2)^2}\sum_{j=1}^{n}\frac{\partial\mu_{y_j|x_j}}{\partial r},$$

$$E\Big\{\Big(\frac{\partial}{\partial\theta_0}\log L\Big)\Big(\frac{\partial}{\partial\theta_1}\log L\Big)\Big\} = \frac{2\varphi^2}{(1-\varphi^2)^2}\sum_{j=1}^{n}\frac{\partial\mu_{y_j|x_j}}{\partial\theta_1},$$

$$E\Big\{\Big(\frac{\partial}{\partial r}\log L\Big)\Big(\frac{\partial}{\partial\theta_1}\log L\Big)\Big\} = \frac{2\varphi^2}{(1-\varphi^2)^2}\sum_{j=1}^{n}\frac{\partial\mu_{y_j|x_j}}{\partial r}\frac{\partial\mu_{y_j|x_j}}{\partial\theta_1},$$

where

$$\frac{\partial\mu_{y_j|x_j}}{\partial r} = \frac{-2\sin(\theta_{x_j}-\theta_1)}{1+2r\cos(\theta_{x_j}-\theta_1)+r^2}, \quad \frac{\partial\mu_{y_j|x_j}}{\partial\theta_1} = \frac{2r\{r+\cos(\theta_{x_j}-\theta_1)\}}{1+2r\cos(\theta_{x_j}-\theta_1)+r^2}.$$

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 $T = -2\log(\max L_0/\max L_1)$, where $\max L_0 = \max_{\theta_0,\varphi} L(\theta_0, \varphi, r = 1, \theta_1 = 0)$, and $\max L_1 = \max_{\theta_0,r,\theta_1,\varphi} L(\theta_0, r, \theta_1, \varphi)$. Under the null hypothesis, T is asymptotically distributed as a chi-square distribution with two degrees 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 used for inference for the proposed model.

5. Examples

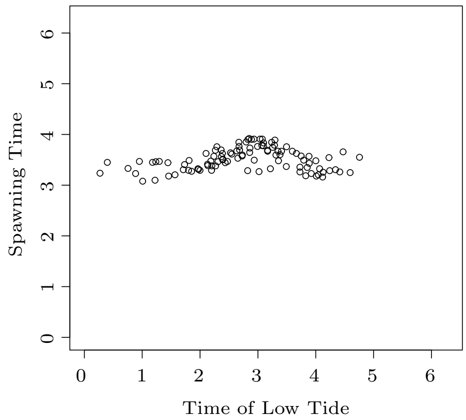
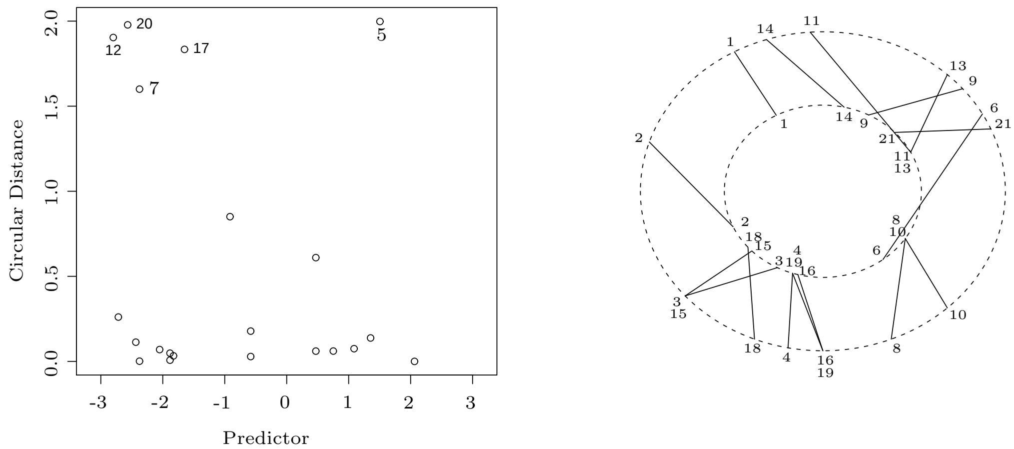


Figure 2. planar plot of the spawning time of certain fish and the time of low tide. Both times are converted into angles $[0, 2\pi)$.

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 24 hours of T_S and T_{LT} to $[0, 2\pi)$. Paired T_S and T_{LT} are thus bivariate circular data, and they are plotted as circles in Figure 2. In the following, we apply model (2.4) to investigate whether and how T_S depends on T_{LT}.

The maximum likelihood estimates of the parameters are $\hat{\theta}_0 = 0.47$, $\hat{r} = 0.95$, $\hat{\theta}_1 = 3.06$ and $\hat{\varphi} = 0.87$. The maximum log-likelihood and AIC of the model are equal to -11.28 and 30.56, respectively. Approximate 90% confidence intervals for θ_0, r, θ_1 and φ are $(-0.11, 1.05)$, $(0.89, 1.00)$, $(2.46, 3.66)$, and $(0.84, 0.90)$, by the Fisher information matrix in Section 4.2. The test of independence for model (2.4) yields the test statistic $T = -2\{(-14.81) - (-11.28)\} = 7.06$. This test is highly significant and the assumption of independence is rejected. Circular distances of all observations lie in $[0, 0.25]$. Here the circular distance is defined by $d(y, \hat{y}) = 1 - \cos(y - \hat{y})$, where y is a response and $\hat{y}$ is a predictor in radians given by $\hat{y} = \hat{\theta}_0 + x - 2\arg\{1 + re^{i(x-\hat{\theta}_1)}\}$.

Example 2. The wind direction at 6 a.m. and 12 noon was measured each day at a weather station in Milwaukee for 21 consecutive days. (Johnson and Wehrly (1977, Table 2)). We use (2.4) for regressing the wind direction at 12 noon on that at 6 a.m.



(a) (b)

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

The maximum likelihood estimates of the parameters are $\hat{\theta}_0 = 1.27$, $\hat{r} = 0.53$, $\hat{\theta}_1 = 2.59$, and $\hat{\varphi} = 0.55$. The maximum log-likelihood and AIC of the model are -32.26 and 72.52, respectively. Approximate 90% confidence intervals for θ_0, r, θ_1 and φ are $(0.91, 1.63)$, $(0.31, 0.74)$, $(2.31, 2.87)$, and $(0.37, 0.73)$. Judging from the AIC, model (2.4) provides a better fit than the Downs and Mardia model, whose AIC is 74.56. The test of independence for (2.4) in Section 4.3 yields the test statistic $T = -2\{(-38.48) - (-32.26)\} = 12.44$. This test is highly significant and the assumption of independence is rejected.

The plot of circular distances is given in Figure 3(a). The observed numbers of outliers are marked on the plot. Apart from five outliers, model (2.4) seems to provide a satisfactory fit to the data. Finally, the predictors and responses, except for the outliers, are plotted by observed numbers in Figure 3(b). The plots on the larger circle refer to the responses, while those on the smaller one are the predictors from (2.4). A short line between the predictor and response means a 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 responses are transformed through the Möbius circle transformation, see Section 2.1.

6. 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 circle transformation and the wrapped Cauchy distribution. As discussed in Section 2.2, the wrapped Cauchy is related to the Möbius circle transformation and thus enables us to obtain a number of desirable properties for our model. As for extensions, our regression model could provide some topics to related fields. For example, the related bivariate circular distribution, which is briefly discussed in Section 3, could be a possible field for further research. It could be also interesting to investigate the properties of the regression model which has the angular error proposed by Jones and Pewsey (2005) instead of the wrapped Cauchy in (2.4). Their model includes the wrapped Cauchy and von Mises as special cases, and might be used to discriminate in applications between these two distributions.

Acknowledgement

We are grateful to the Associate Editor and two referees for their valuable comments on our paper. Our thanks also go to Dr. Nicholas I. Fisher and Dr. Peter J. Thomson who helped us improve the quality of the first draft of our paper. This work of S. Kato and K. Shimizu 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.

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

Jones, M. C. (2004). The Möbius distribution on the disc. *Ann. Inst. Statist. Math.* **56**, 733-742.

Jones, M. C. and Pewsey, A. (2005). A family of symmetric distributions on the circle. *J. Amer. Statist. Assoc.* **100**, 1422-1428.

Kent, J. T. and Tyler, D. E. (1988). Maximum likelihood estimation for the wrapped Cauchy distribution. *J. Appl. Stat.* **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. Stat.* **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. Amer. 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. *Comm. Statist. Theory Methods* **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. Probab. Lett.* **12**, 373-378.

Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan.

E-mail: skato@ism.ac.jp

Department of Mathematics, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama 223-8522, Japan.

E-mail: shimizu@stat.math.keio.ac.jp

Institute of Statistical Science, Academia Sinica, Taipei 115, Taiwan, R.O.C.

E-mail: gshieh@stat.sinica.edu.tw

(Received April 2006; accepted October 2006)