ON FINDING THE LOCATION OF A

SIGNAL: A BAYESIAN ANALYSIS

Peter Guttrop, Richard A. Lockhart

Tech. Report No. 97
ON FINDING THE LOCATION OF A SIGNAL:
A BAYESIAN ANALYSIS

by

P. Guttormp, University of Washington
Richard A. Lockhart, Simon Fraser University

Tech. Report # 97
June 1986

Prepared in part under support to SIMS from the United States Environmental Protection Agency
Department of Statistics, University of British Columbia, 2021 West Mall,
Vancouver, B.C. V6T 1W5
On finding the location of a signal: a Bayesian analysis

Peter Guttrop
Department of Statistics
University of Washington
Seattle, WA
USA

Richard A. Lockhart
Department of Mathematics and Statistics
Simon Fraser University
Burnaby, B.C.
Canada
ABSTRACT

We study the problem of determining the location of an emergency transmitter from directional observations at fixed stations. In doing so, we develop conjugate prior distributions for the von Mises distribution for one or both parameters unknown. We develop a sensitivity analysis for the bearings which can be used to eliminate wild readings.

**Key words**: Directional data, von Mises distribution, conjugate prior, emergency locator transmitter.

*Acknowledgements*: Partial support from the United States Environmental Protection Agency under a cooperative research agreement with SIMS. Part of this work was done while the first author was visiting the University of British Columbia.
1. Introduction

Every year there are several thousand general aviation accidents in the United States. Many of them result in fatalities. About 200,000 civilian airplanes in the US are each equipped with an emergency locator transmitter (ELT). These transmitters are designed to transmit a distress signal from a downed aircraft, immediately and automatically upon impact. Suppose that we have \( n \) receivers at specified locations \( x_1, \ldots, x_n \), and an emergency transmitter is activated somewhere within receptive distance of these locations. Each receiver determines the direction \( \theta_i \) of highest intensity. Lenth (1981) develops maximum likelihood and robust estimates of the location of the transmitter. We find a Bayesian analysis more natural, since one wants relative probabilities of the ELT being in different spots.

Bayes methods for directional data have received little attention in the literature. In fact, Kikuchi's (1982) collection of directional data abstracts from the decade following the publication of Mardia (1972) only mentions one paper on Bayes methods: Mardia and El-Atoum (1976), in which they discuss Bayes estimation for the von Mises-Fisher distribution with known concentration parameter. In section 2 of this note we develop conjugate prior distributions for the von Mises distribution with unknown concentration and either known or unknown location. In section 3 we give a Bayesian analysis of the ELT problem, utilizing some of the tools developed in section 2. We obtain a posterior density for the location of the emergency transmitter. This can be thought of as a probability map to guide the search effort. Section 4 of the paper contains some discussion in the form of a series of remarks.

2. Bayes inference for the von Mises distribution

Let \( y_1, \ldots, y_n \) be observations (unit vectors) of a random variable \( Y \) with a von Mises distribution on the circle (von Mises, 1918). We will write the density of \( Y \) with respect to Lebesgue measure on the circle in canonical exponential family form as

\[
f(y | \tau) = \frac{1}{2\pi I_0(|\tau|)} \exp(\tau' y), \quad |y| = 1,
\]  

where \( I_0 \) is a modified Bessel function of the first kind and order zero, \( \tau \) is the canonical parameter vector,
and $|\tau|=(\tau')^{1/2}$. In terms of the conventional parametrization with a mean direction $\theta$ and a concentration parameter $\kappa$ we can write $\tau'=\kappa(\cos\theta, \sin\theta)$. The canonical parameter space is $\mathbb{R}^2$. In what follows we will develop conjugate prior distributions for this family for different combinations of known and unknown parameters.

2.1. Known concentration, unknown direction (Mardia and El-Atoum, 1976)

A prior distribution, conjugate to (1), is given by

$$f(\theta) = I_0(c \kappa)^{-1} \exp(\kappa y_0' \tau).$$

(2)

Let $y_0' = c(\cos\theta_0, \sin\theta_0)$. The posterior distribution after observing a sample $y=(y_1, \ldots, y_n)$ is a von Mises distribution with location parameter $\theta_n = \frac{\sum y_i}{\sum y_i}$ and concentration parameter $\kappa_n = \kappa \sum y_i$. The prior parameters can be thought of as representing $c$ observations all in the direction $\theta_0$. The Bayes estimate of $\theta$ with respect to squared error loss on the unit circle is $\cos \theta_n$. In the limit, as $c \to 0$, the prior converges to the uniform prior on the circle, which is the natural way of expressing vague prior knowledge.

2.2. Unknown concentration, known direction

There is no lost generality in assuming that the known direction is zero. A conjugate prior is then given by

$$f(\kappa) \propto I_0(\kappa)^{-c} \exp(\kappa R_0).$$

(3)

The constant of proportionality is $B_0^{-1}(R_0, c)$, where

$$B_\alpha(\beta, \gamma) = \int_0^\infty x^{\alpha-1} I_0(x) \exp(-\beta x) dx, \quad a < b.$$

Let $R_n = R_0 + \sum \cos \theta_i$. Then the posterior density can be written

$$f(\kappa | y) = \frac{\exp(\kappa R_n)}{I_0(\kappa)^{c+n} B_0(R_n, c+n)}.$$
The prior parameter $R_0$ can be thought of as the component on the x-axis (i.e., the known direction) of the resultant of $c$ observations.

A simple approximate formula for $B_\alpha$ is based on the approximation

$$I_0(x) - (2\pi x)^{-\frac{\gamma}{2}} \exp(x), \quad x \to \infty$$

(Abramowitz and Stegun, 1965, equation 9.7.1), namely

$$B_\alpha(\beta, \gamma) = (2\pi)^{\frac{\gamma}{2}} \Gamma\left(\frac{\gamma}{2} + \alpha + 1\right)(\gamma - \beta)^{-\frac{\gamma}{2} + \alpha + 1}.$$  

The Bayes estimate with respect to squared error loss is

$$E(\kappa | y) = \frac{B(\kappa, c + n)}{B_0(R_n, c + n)}.$$  

The approximation (4) yields $E(\kappa | y) = (c + n + 2)/(c + n - R_n)$. This is reminiscent of the precision parameter estimate $n/2(n - R_n^*)$, where $R_n^*$ is the length of the resultant of $y_1, \ldots, y_n$, obtained using the method of moments in Mardia (1972).

Another popular Bayes estimate is the posterior mode, which is seen to satisfy the equation

$$A(\kappa) = \frac{R_n}{c + n},$$

where $A(\kappa) = I_1(x)/I_0(x) = -\frac{d}{dx} \log I_0(x)$.

The Fisher information about $\kappa$ in an observation $y$ is $A'(\kappa)$. The Jeffreys prior (Jeffreys, 1961, ch. 3.10) is taken proportional to the square root of the absolute value of the Fisher information. The approximation (4) shows that for large values of $\kappa$, the Jeffreys prior is proportional to $\kappa^{-1}$. This prior is often used, particularly in one parameter situations, to describe vague knowledge. If one attempts to use (3) to express vague prior information by setting $R_0 = \delta c$ for some $0 < \delta < 1$, and the let $c \to 0$, one obtains the uniform prior on the positive line. This is usually not a reasonable description of vague knowledge, since it represents a situation when large values of $\kappa$ are thought rather likely a priori.
2.3. Both parameters unknown

In this case a conjugate prior for $\tau$ has density

$$f(\tau) \propto I_0(|\tau|)^c \exp(\tau y_0),$$

(5)

where $y_0 = R_0(\cos \theta_0, \sin \theta_0)$ and $c$ are prior parameters. Let

$$H_0(\beta, \gamma) = \int_0^\infty x^\beta I_0(\beta x) I_0(x)^\gamma dx.$$  

The normalizing constant for $f(\tau)$ is $H_1(R_0, c)^{-1}$. The posterior density of $\tau$, given the sample $y$, is

$$f(\tau | y) = \frac{[I_0(|\tau|)^c H_1(R_0, c + n)]^{-1} \exp(\tau R_0(\cos \theta_n, \sin \theta_n))}{H_0(\beta, \gamma)}$$

(6)

where $R_n$ is the length of the resultant of $y_0, y_1, \ldots, y_n$ and $\theta_n$ is the angle of the resultant.

The Bayes estimates of $\tau$ with respect to squared error loss are now readily computed.

$$E(\tau_1 | y) = \int_0^{2\pi} \int_0^\infty \frac{\kappa^2 \cos \theta \exp(\kappa R_n \cos(\theta - \theta_n))}{H_1(R_n, c + n) I_0(\kappa)^c + n} d\theta d\kappa$$

$$= \cos \theta_n \frac{d}{dR_n} \log H_1(R_n, c + n).$$

Likewise

$$E(\tau_2 | y) = \sin \theta_n \frac{d}{dR_n} \log H_1(R_n, c + n).$$

The Bayes estimate of $\cos \theta$ with respect to squared error loss on the unit circle is seen to be $\cos \theta_n$ as in section 2.1. There is no convenient closed form for the Bayes estimate of $\kappa$ with respect to squared error loss, but for large values of $\kappa$ a simple approximation to the Bayes estimate of $\kappa$ is obtained from (4) as in the previous section:

$$E(\kappa | y) = \frac{c + n + 3}{2(c + n - R_n)}.$$

Straightforward calculations show that the posterior modal point solves
\[ A(\mid \tau \mid) = R_n / (n + c) \]
\[
\arg \tau = \theta_n.
\]

From Mardia (1972) we see that the posterior mode estimates of \(|\tau| = \kappa\) and \(\arg \tau = \theta\) are just the maximum likelihood estimates from the sample \(y\), augmented by \(c\) observations with resultant \(y_0\). This provides a natural interpretation of the prior parameters \(c\) and \(y_0\).

Marginalizing (6) and changing variables, we find the posterior densities of \(\kappa\) and \(\theta\) separately as

\[
f(\kappa \mid y) = \frac{\kappa I_0(\kappa R_n)}{I_0(\kappa)^{c+n} H_1(R_n, c + n)}
\]

and

\[
f(\theta \mid y) = \frac{\kappa \cos(\theta - \theta_n)}{I_0(\kappa)^{c+n} H_1(R_n, c + n)}
\]

The Fisher information about \(\tau\) in an observation from the density (1) is

\[
A(\mid \tau \mid) \frac{\text{id}_2}{\mid \tau \mid} - \frac{\tau \tau'}{\mid \tau \mid^2 (1 - A^2(\mid \tau \mid))}
\]

where \(\text{id}_2\) is the 2x2 identity matrix. The determinant of the Fisher information is

\[
- \frac{A(\mid \tau \mid)}{\mid \tau \mid} A'(\mid \tau \mid).
\]

The Jeffreys prior is proportional to the square root of the absolute value of (7). Mardia and El-Atoum (1976) give the Jeffreys prior for the \((\kappa, \theta)\)-parametrization. The canonical parametrization simplifies the form of the prior. If we try to formalize the idea of no prior knowledge using (5) by letting \(c\) and \(R_0\) approach zero with \(R_0 = \delta c\), \(0 < \delta < 1\), we obtain the (improper) uniform prior on \(\mathbb{R}^2\). In the \((\kappa, \theta)\)-space this corresponds to a uniform prior on \(\theta\) given \(\kappa\), as prescribed by the Jeffreys prior, but a prior on \(\kappa\) proportional to \(\kappa\). Thus \(\kappa\) is thought even larger \textit{a priori} than in the case of known \(\theta\).
3. Locating emergency transmitters

When an airplane crashes, an emergency locator transmitter is activated. The nature of the signal from this device, at 121.5 Mc/s, is such that severe multipath distortion may occur in mountainous terrain. It is naturally of high importance to locate the crash site as rapidly as possible. There are joint programs between the United States and the Soviet Union for satellite relay systems to determine the location (Waldrop, 1984), but here we will concentrate on the more conventional system with fixed stations using radio direction finding equipment, combined with visual airplane search. The accuracy of these searches depend upon the ability of the pilot to fly a prescribed search pattern, and his ability to recognize the terrain beneath him to direct the ground crew to the proper area. In periods of bad weather, when accidents are more likely to occur, the search and rescue ground crew may have to depend entirely on the radio direction finding technique.

An early reference in the statistical literature is Daniels (1951), who reports on his war experience of position finding. He applies standard theory of errors, in assuming that observation errors displaces the bearing line parallel to itself, and computes a least squares estimate of location. He also develops confidence bands, using normally distributed errors. Gething (1978, Ch. 14) points out that typical bearing errors, apart from containing a fraction of "wild" observations, also are leptokurtic. From a statistical point of view, it appears more natural to assign a directional distribution to the measurement errors.

There are several sources of error in direction finding data. These are discussed e. g. in Gething (op. cit.). Objects near the receptor site, such as wire fences, railways, power lines, buildings etc. can introduce severe irregularities in the readings. Propagation errors are introduced by orographic features for ground paths, and by the variability of the ionosphere for sky waves. The latter are less important for VHF signals.

We assume that bearings are taken independently from stations at sites \( x_1, \ldots, x_n \). The true location of the target (the crash site) is \( x \), and the bearing from the i'th station is \( \theta_i \). Let \( \mu_i \) be the true direction from station \( i \) to the target, so that

\[
\tan(\mu_i) = \tan(\mu_i(x)) = \frac{x(2) - x(1)}{x(1) - x(1)},
\]
where \( x_i(j) \) is the \( j \)'th coordinate of \( x_i \), and let \( \kappa \) denote the precision of the receiving instrument. We assume that, given \( x \) and \( \kappa \), the \( \theta_i \) are independent random variables with von Mises densities

\[
f(\theta_i|x,\kappa) = \frac{\exp(\kappa \cos(\theta_i - \mu_i))}{2\pi I_0(\kappa)}
\]  

We will assume that the device accuracy parameter \( \kappa \) has the prior (3). Marginalizing (8) with respect to the prior density of \( \kappa \) we get, with \( \theta'=(\theta_1, \ldots, \theta_n) \), that

\[
f(\theta|x) = \int f(\theta|x,\kappa)f(\kappa)d\kappa = (2\pi)^{-n} \frac{B(R_n(x),c+n)}{B(R_0,c)},
\]

where \( R_n(x) = R_0 + \sum \cos(\theta_i - \mu_i(x)) \). The posterior density of \( x \) given \( \theta \) is proportional to \( f(\theta|x)f(x) \), where \( f(x) \) is the prior density of \( x \). It may often be reasonable to use a flat prior over the region of receptability. In other situations where likely spots for accidents are known, this should of course be reflected in the choice of \( f(x) \).

We apply this method to the data from Lenth (1981), reproduced in Table 1. Figure 1 shows the stations and their bearings.

*** Table 1 about here ***

We use three different choices of prior parameters. The first is based on the actual errors (knowing the true location) in Lenth's data, and can be thought of as corresponding to prior information from a calibration experiment with seven stations and a known location. We fit \( \kappa \) by maximum likelihood, yielding \( \kappa = 307.1 \) with a standard error estimated at 164.0. This corresponds to an error standard deviation of 3.3°, and the corresponding choice of prior parameters is \((R_0,c)=(4.9836,4.995)\). Figure 2 shows the posterior density, using a flat \( f(x) \). The contours are chosen on a logarithmic scale.

Gething (1978) shows some data for HF direction finding, using wide aperture direction finding stations (usually more precise than those used in ELT direction finding), and calculates a trimmed standard deviation of 1.7°. Bowen (1955), studying systems more like those employed in this situation, reports a standard
deviation of 2.9°. Gökeri (1965) reports an experiment where fairly distant readings yielded trimmed variances ranging from 2.4° to 5.0°. For the higher values of variance, the trimming was probably too severe. Thus, the value used above does not appear unreasonable. In order to get some idea of the importance of the prior on the posterior map, we also tried one high and one low precision prior. The high precision values had a prior mean of \( \kappa \) corresponding to an error standard deviation of 1°, with a prior standard deviation about 0.3°, corresponding to \((R_{0}, c) = (4.99893, 5.0)\). The low precision values had an expected precision corresponding to a bearing standard deviation of 8°, with prior standard deviation of 1.6°. The parameter values are \((R_{0}, c) = (9.8824, 10)\). Figure 3 shows the 99% probability contours for the three different priors. These shrink with increasing precision, as would be expected. The contours cover the true value, although the data indicate a probable location southwest of the target.

In order to study the sensitivity of the procedure to unusual or incorrect observations, we added an eighth station with an incorrect bearing, off by about 20°. Lenth (1981) used an error of about 90°. As can be seen from Figure 1, such an error is immediately visible from the plot, whereas our choice is not completely inconsistent with the other observations. The effect is shown in Figure 4, depicting the 99% probability contour of the posterior with and without the additional observation. We used the prior parameters corresponding to intermediate precision, as in Figure 2. The size of the probability contour for the posterior including the additional observation indicates that the precision of the data has been overestimated by the prior. The contour has been shifted further to the southwest, although the target still remains inside the contour.

Box (1983) proposes the use of predictive distributions to judge the quality (or influence) of a datum used in the analysis. The predictive density of a model, parametrized by \( \eta \), can be written

\[
p(z) = \int p(z | \eta)p(\eta) d \eta
\]

where \( p(\eta) \) is the prior. If the observed datum \( z_o \) is extreme with respect to \( p(z) \), the model is considered doubtful. By looking at each datum component separately, one can assess the quality of individual components. We combine this influence measure with the idea of crossvalidation. Let \( \theta^{-i} \) be the data vector,
leaving out the \( i \) th observation. The crossvalidated predictive likelihood of the \( i \) th observation is

\[
f (\theta \mid \theta^{-i}) = \frac{f (\theta)}{f (\theta^{-i})} = \frac{\int f (\theta \mid x) f (x) dx}{\int f (\theta^{-i} \mid x) f (x) dx} = \frac{\int B (R_n (x, \theta), c + n) f (x) dx}{\int B (R_{n-1} (x, \theta^{-i}), c + n - 1) f (x) dx}
\]

where, for clarity, we make explicit the dependence of \( R_n (x) \) on \( \theta \). In Table 2 we give the predictive likelihoods for the original data as well as for the data augmented with the errant reading used above.

*** Table 2 about here ***

The predictive likelihood for the errant reading is an order of magnitude smaller than that of any other station, using the data-based parameter values \((c, R) = (4.9836, 4.995)\).

The computations in this paper were performed on a VAX 11/750 using IMSL routines for numerical integration (DCADRE) and computation of Bessel functions (MMBSI0 and MMBSI1). Although there were no major numerical problem in this analysis, the computation time for the sensitivity analysis is rather long. As an alternative, we used the approximation based on (4), which yields a posterior density (using a flat prior on \( x \)) proportional to \((c + n - R_n (x))^{-(n + c + 2)/2}\). The computation of the posterior map using this approximation is quite fast. The quality of the approximation is high. For the intermediate precision prior, the absolute approximation error is 0.0065 at the maximum density value 6.65. Usually the approximation is too high by a small amount. The maximum absolute relative error is 2.67, obtained where the density is \(9.0 \times 10^{-17}\). The corresponding approximation to the formula (9) for the sensitivity analysis is

\[
f (\theta \mid \theta^{-i}) = (2\pi)^{\frac{1}{2}} \frac{\Gamma \left( \frac{c + n + 2}{2} \right) \int (c + n - R_n (x, \theta))^{-(c + n + 2)/2} dx}{\Gamma \left( \frac{c + n + 1}{2} \right) \int (c + n - 1 - R_{n-1} (x, \theta^{-i}))^{-(c + n + 1)/2} dx}
\]

The computation of the approximate formula, using summation over a coarse grid (square side 0.25), takes less than a minute of CPU time on the VAX, whereas the exact formula, with a numerical integration for each
function evaluation, takes about 4 hours of CPU time. The approximation is good, as seen in Table 2. The adequacy of summing over a coarse grid as a method of double integration was tested by using the IMSL double integration routine DBLIM. The results agreed to two decimal places, which is better accuracy than that of the approximation.

Using the approximation (10) we investigated what size error can be distinguished using the sensitivity analysis. Letting the error in the observation for station 6 vary from -90° to 90°, we show in Figure 5 the ratio of smallest to second smallest predictive likelihood. A crude decision rule, disallowing a reading if this ratio is smaller than 0.1, rejects all readings on station 6 except for an error between -5° and 10°. The ratio between the second smallest and the largest predictive likelihood varied between 0.11 and 0.35, and the maximum predictive likelihood varied from 7.9 (corresponding to an error of -90°) to 54.8 (corresponding to 90°).

4. Discussion

Remark 1: In some situations it may be useful to introduce a bias parameter associated with the reception stations. The data density (8) would be replaced by

\[ f(\theta_j | x, \kappa, \beta) = \frac{\exp(\kappa \cos (\theta_j - \mu_i - \beta))}{2\pi l_0(\kappa)}. \]

We assume that the device accuracy parameters \( \kappa \) have the prior (5). The bias parameter may be thought of as representing peculiar atmospheric conditions. In such circumstances it may be natural to choose \( \theta_0 = 0 \). This particular choice of prior means that large biases are rather more likely when the precision is small than when it is large. Calculations similar to those in section 3 show that the posterior density of \( x \), given \( \theta \), is proportional to \( H_1(|\Delta|, c + n) f(x) \). For small values of precision, this tends to produce banana-shaped ridges in the posterior density. The effect of the bias parameter disappears as the precision increases.

Remark 2: The prior parameters in the case of known precision can be given a different interpretation from that in section 2.1. Define \( \eta = \log I_0(c \kappa) - \log I_0(\kappa) \). Then the prior (2) becomes

\[ f(\theta) = l_0^{-\eta}(\kappa) \exp(\kappa y_0 \tau) \]
and the interpretation is that of \( \eta \) observations with resultant \( y_0 \). For large values of \( \kappa \), the approximation (4) shows that \( \eta = c \).

**Remark 3:** The standard statistical technique used in radio direction finding appears to be least squares, based on the assumption of normally distributed errors, as in Daniels (1951). Assume that \( \theta_i - \mu_i \), given \( \kappa \) and \( x \), are normally distributed with mean 0 and variance \( \kappa^{-1} \). Using the conjugate gamma prior on \( \kappa \), standard computations show that

\[
f(x | \theta) \propto \frac{f(x)}{d(x)^{c+n+1}}
\]

(9)

where \( d(x) = 2(c - R_0) + \sum(\theta_i - \mu_i(x))^2 \).

The analysis based on normality can be thought of as introducing two steps of approximation to the posterior density

\[
f(x | \theta) \propto f(x) B_0(R_n(x), c+n).
\]

First, using (4) to approximate \( B_0(R_n, n+c) \) we get

\[
f(x | \theta) \propto \frac{f(x)}{(c+n-R_n(x))^{c+n+1}}.
\]

Second, the quantity \( R_n(x) \) is approximated, using a Taylor expansion, by

\[
R_n(x) = R_0 + n - \frac{1}{2} \sum(\theta_i - \mu_i)^2,
\]

yielding the posterior density (9).

Two points should be made about this approximation. First, we can write

\[
R_n(x) = R_0 + \sum \cos(\theta_i - \mu_i) = R_0 + \sum \frac{(\cos \theta_i)(x_i(1) - x_1) + (\sin \theta_i)(x_i(2) - x_2)}{p(x, x_i)}
\]

where \( p(x, x_i) \) is the Euclidean distance between the two points. The evaluation of this quantity is considerably simpler than the evaluation of \( \sum(\theta_i - \mu_i(x))^2 \), which involves computing arctangents in each
term. Second, although the Taylor approximation of the cosine function is excellent for small values of the argument, it is not so good for large values. The latter can be expected fairly frequently, especially for data in mountainous terrain. Therefore the normal model is not expected to perform well in the data validation part of the analysis.

**Remark 4:** It may sometimes be important to update the distribution of $\kappa$, particularly in early stages of calibration of a new station. We compute

$$f(\kappa | \theta) \propto \frac{f(\kappa)}{\int_0^\infty f(\kappa) \exp\{\kappa(R_n(x) - R_0)\} f(x) dx}.$$  

In particular, one may be interested in the posterior mean, which is seen to be

$$E(\kappa | \theta) = \frac{\int B_1(R_n(x), n+c)f(x) dx}{\int B_0(R_n(x), n+c)f(x) dx}.$$  

The integral in the denominator has already been computed during the sensitivity analysis.

**Remark 5:** A more general model than that envisioned in section 3 would drop the prior independence assumption between $\kappa$ and $x$. Although we have not been able to produce exact solutions in reasonable form, the large $\kappa$ approximation allows us to analyze the problem. Assuming that $\kappa$ given $x$ is inversely proportional to a power of the distance, namely $\kappa_i = \kappa_0 \rho(x_i, x)^{-\alpha}$, and assigning the prior (3) to $\kappa_0$, we get an approximate posterior density proportional to

$$f(x | \theta) \propto f(x) R_n(x, x)^{-c+n+2})^2 \prod_i \rho(x_i, x)^{-\alpha/2}$$

where $R_n(x, x) = \sum_i \rho(x_i, x)^{-\alpha} [1 - \cos(\theta_i - \mu_i(x))]$.

**References**


Table 1. Data from Lenth (1981)

<table>
<thead>
<tr>
<th>No.</th>
<th>Location</th>
<th>Compass reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.0,3.2</td>
<td>234°</td>
</tr>
<tr>
<td>2</td>
<td>9.4,5.5</td>
<td>215°</td>
</tr>
<tr>
<td>3</td>
<td>9.95,9.05</td>
<td>196°</td>
</tr>
<tr>
<td>4</td>
<td>8.7,8.7</td>
<td>193°</td>
</tr>
<tr>
<td>5</td>
<td>8.07,9.7</td>
<td>188°</td>
</tr>
<tr>
<td>7</td>
<td>4.4,8.9</td>
<td>160°</td>
</tr>
<tr>
<td>8</td>
<td>1.85,5.25</td>
<td>118°</td>
</tr>
<tr>
<td>6</td>
<td>5.6,9.0</td>
<td>(180°)*</td>
</tr>
</tbody>
</table>

*This station was only used for the sensitivity analysis. There was no actual reading at this station. The error is about 20°. Lenth (1981) used an observation of 250° for this station.

Table 2. Predictive likelihoods

<table>
<thead>
<tr>
<th>Station no.</th>
<th>Likelihood excluding 6</th>
<th>Approximation using (10)</th>
<th>Likelihood including 6</th>
<th>Approximation using (10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18.0</td>
<td>18.1</td>
<td>19.8</td>
<td>19.1</td>
</tr>
<tr>
<td>2</td>
<td>27.1</td>
<td>27.1</td>
<td>34.9</td>
<td>33.7</td>
</tr>
<tr>
<td>3</td>
<td>14.3</td>
<td>14.3</td>
<td>14.2</td>
<td>13.7</td>
</tr>
<tr>
<td>4</td>
<td>51.8</td>
<td>51.9</td>
<td>38.9</td>
<td>37.5</td>
</tr>
<tr>
<td>5</td>
<td>46.2</td>
<td>46.3</td>
<td>40.5</td>
<td>39.1</td>
</tr>
<tr>
<td>7</td>
<td>39.6</td>
<td>39.7</td>
<td>38.4</td>
<td>37.0</td>
</tr>
<tr>
<td>8</td>
<td>17.3</td>
<td>17.4</td>
<td>5.9</td>
<td>5.7</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>0.60</td>
<td>0.58</td>
</tr>
</tbody>
</table>
Figure captions

Figure 1. Locations of the stations in Lenth (1981). The lines represent observations. The bearing at station 6 is used for sensitivity analysis only. The asterisk represents the target.

Figure 2. Log posterior probability map for the prior parameters \((R_0, c) = (4.9836,4.995)\). The asterisk represents the target.

Figure 3. 99% posterior density contours for three different priors. The inside short dash curve, corresponds to the high precision prior \((R_0, c) = (4.9989,5.0)\). The middle solid curve uses the prior \((R_0, c) = (4.9836,4.995)\). The outside long dash curve stands for the low precision prior \((R_0, c) = (9.8824,10.0)\). The asterisk represents the target.

Figure 4. 99% posterior density contours for the intermediate precision prior \((R_0, c) = (4.9836,4.995)\). The solid line corresponds to the original data, while the dashed line is the posterior density after adding a reading with 20° error at station 6. The asterisk represents the target.

Figure 5. The ratio of smallest to second smallest predictive likelihood as a function of the error at station 6.