Comput Stat
DOI 10.1007/s00180-012-0307-3

ORIGINAL PAPER

Monte Carlo maximum likelihood circle fitting using circular density functions

Ulric J. Lund

Received: 29 March 2011 / Accepted: 24 January 2012

Abstract Finding the "best-fitting" circle to describe a set of points in two dimensions is discussed in terms of maximum likelihood estimation. Several combinations of distributions are proposed to describe the stochastic nature of points in the plane, as the points are considered to have a common, typically unknown center, a random radius, and random angular orientation. A Monte Carlo search algorithm over part of the parameter space is suggested for finding the maximum likelihood parameter estimates. Examples are presented, and comparisons are drawn between circles fit by this proposed method, least squares, and other maximum likelihood methods found in the literature.

Keywords Circle fitting · Maximum likelihood · Least squares · von Mises distribution · Minimum-width annulus · Monte Carlo

1 Introduction

The optimization problem of finding the "best-fitting" circle to a set of points in two dimensions arises in many scientific applications. Examples can be found in fields such as mechanical engineering (Ahn et al. 2004), biometric identification (Ertürk 2006), forestry (Thomas et al. 2007), electrical engineering (Janóczki and László 2010), and archaeology (Rorres and Romano 1997; Schaffrin and Snow 2010). Typical formulations of the circle fitting problem are in terms of a least squares or maximum likelihood approach. In either case, it is of interest to estimate the unknown parameters that describe an "optimal" circle's center location and radius, that along with random error could have given rise to a set of observed (X, Y) coordinate data.

U. J. Lund (✉)
Department of Statistics, California Polytechnic State University, San Luis Obispo, CA, USA
e-mail: ulund@calpoly.edu

Published online: 09 February 2012

Several references regarding the statistical properties of circle parameter estimates are as follows. Chan (1965) investigated the consistency of estimators. Chan and Thomas (1995) computed the Cramer-Rao lower bounds of circle parameter estimates under Gaussian noise. Kasa (1976) and Berman and Culpin (1986) looked into some characteristics of least squares estimators of the circle parameters.

Some computational aspects of obtaining least squares estimates is discussed in Coope (1993). Umbach and Jones (2003) showed how various least squares methods compare in terms of their sensitivity to measurement error. Chernov and Lesort (2005) investigate the instability of least squares estimators. Schaffrin and Snow (2010) introduce a Tykhonov regularized total least squares approach.

Mardia and Holmes (1980) discussed possible maximum likelihood models for elliptical data. Anderson (1981) discussed the asymptotic behavior of some maximum likelihood estimators, and proposed using an approximation to the ratio of modified Bessel functions in the estimation process. Chan et al. (2005) showed that an approximate maximum likelihood algorithm can yield estimates close to the Cramer-Rao lower bounds, and Chernov and Sapirstein (2008) looked at maximum likelihood estimation in the presence of correlated noise. A comparison of the statistical properties of some maximum likelihood and least squares estimators is given in Berman and Culpin (1986) and Berman (1989).

The maximum likelihood approaches in the existing literature typically assume a bivariate Gaussian distribution on the (X, Y) coordinates. Only three of the papers above considered circular distributions in the stochastic models for the data (Mardia and Holmes 1980; Berman and Culpin 1986; Berman 1989), and of these only Berman (1989) considered the case when the data is not uniformly distributed around the entire circle.

My article will present a variety of probability models that can be useful in describing the stochastic nature of data that is randomly scattered around a circle or part of a circle. I will make use of circular density functions that can provide deeper insight into the nature of the data than previously considered maximum likelihood models can. I will also present a Monte Carlo estimation technique for arriving at the maximum likelihood estimates of the circle parameters. The considerable progress in computational speed that has been seen since the appearance of the early articles on maximum likelihood circle fitting, and the advent of the bootstrap, can now be leveraged to update the status quo of maximum likelihood circle fitting.

2 Model formulation and estimation

2.1 Model assumptions

To begin the discussion, consider n observations of random cartesian coordinates in the plane, (X, Y); the data being denoted by $(x_1, y_1), \ldots, (x_n, y_n)$. Assume that each random cartesian coordinate arises from a common center (α, β), with a random radius

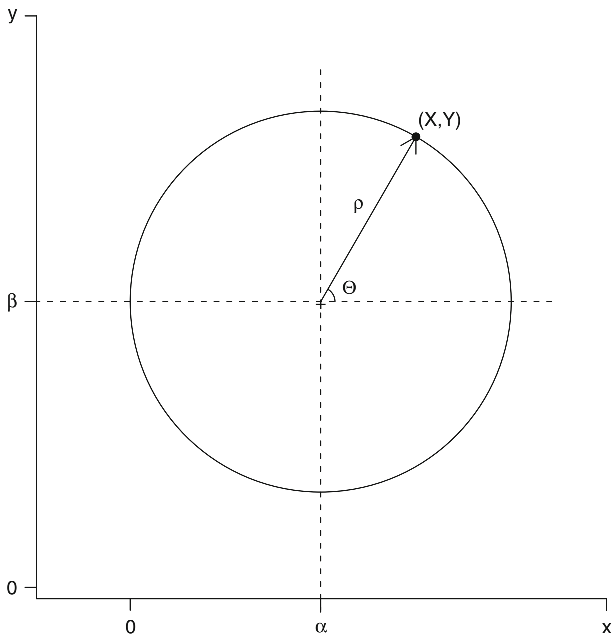

Fig. 1 Cartesian to polar coordinate transformation

or distance from the center ρ, at a random orientation or angle of rotation from the origin Θ radians (see Fig. 1). That is,

$$X = \alpha + \rho \cos \Theta$$
$$Y = \beta + \rho \sin \Theta$$

The data consists of observations of (X, Y). However, the distributional assumptions will be placed on a point's random radius from the center of the circle and its random angular orientation, both of which are unobservable if the center is unknown. This type of model has been called the radial model by Berman and Culpin (1986). Let $f_{\rho,\Theta}(r, \theta|\Psi)$ be the joint denisty for ρ and Θ, where this joint distribution has a vector of parameters given by Ψ. Noting that

$$\Theta = \tan^{-1}\left(\frac{Y-\beta}{X-\alpha}\right) \quad \text{and} \quad \rho = \sqrt{(X-\alpha)^2 + (Y-\beta)^2}, \tag{1}$$

it can be shown that the joint density function of X and Y is given by

$$f_{X,Y}(x, y|\Psi) = \frac{f_{\rho,\Theta}\left(\sqrt{(x-\alpha)^2 + (y-\beta)^2}, \tan^{-1}\left(\frac{y-\beta}{x-\alpha}\right)|\Psi\right)}{\sqrt{(x-\alpha)^2 + (y-\beta)^2}} \tag{2}$$

Table 1 Probability models considered for angle Θ and radius ρ

Case	Probability models	
	Θ	ρ
1	Uniform $(0, 2\pi)$	Uniform (l_ρ, u_ρ)
2		Normal (μ_ρ, σ^2)
3		Gamma (r, c)
4	Uniform (l_Θ, u_Θ)	Uniform (l_ρ, u_ρ)
5		Normal (μ_ρ, σ^2)
6		Gamma (r, c)
7	von Mises (μ_Θ, κ)	Uniform (l_ρ, u_ρ)
8		Normal (μ_ρ, σ^2)
9		Gamma (r, c)

This can be simplified somewhat if it is assumed that the radius and angular rotation are independent, in which case the numerator in (2) can be split into the product of the marginal densities of ρ and Θ.

Using (2), the likelihood function for Ψ could be found in terms of the observable data (x, y). The likelihood function can get quite complex, given that the unknown center's parameters α and β are intertwined in the joint density function. Finding the maximum likelihood estimates (MLE's) for Ψ, α and β is conceptually a straight forward task. But doing so in practice is another matter, given the complexity of the likelihood function.

In this paper, I will assume that ρ and Θ are independent. This is certainly a strong assumption, but one that is almost universally made in the current literature. Chernov and Sapirstein (2008) consider circle fitting when the noise is correlated among observations, which is a different kind of dependence than the one I am assuming. I will suggest a way of testing the assumption of independence between ρ and θ in Sect. 2.4.

Various probability models could be hypothesized for the data. A circular density function is needed to model the angular rotation Θ, and a linear density function is needed for the random radius ρ. Table 1 lists a variety of cases that could be considered in practice. Cases 1–3 assume that the data is distributed uniformly around an entire circle. Cases 4–6 also have the data rotationally uniformly distributed, but limited between lower- and upper-bound angles l_Θ and u_Θ. Cases 7–9 assume that the angular rotation follows a von Mises distribution. The von Mises distribution is the circular analog of the normal distribution on the real line. Its density function is symmetric and unimodal around a mean direction μ_Θ with the concentration (dispersion) of the distribution determined by another parameter κ. The von Mises distribution is a commonly used circular density function, making it a natural choice for modeling the angular rotation Θ. The uniform, normal, and gamma distributions may be reasonable choices for probability models for the random radius.

Probability models other than those listed in Table 1 are certainly possible. For example a truncated normal distribution for the radius may be considered a viable option, truncating at a lower bound of 0, to avoid the theoretical possibility of negative radii. The circle fitting method illustrated later in this article can be tailored to such other models as well.

The cases in Table 1 provide a much richer model for circle data, compared to what is assumed under the traditional least squares formulation of the problem. Actually, the least squares formulation does not require any distributional assumptions on the data. However, the characteristics of the least squares estimators have been explored under the model that X and Y are observed with bivariate Gaussian random error. This type of model, and the least squares formulation in general, does not allow one to estimate an arc in which the data falls, while Cases 3 and 4 do; the least squares formulation does not allow one to estimate an average angular orientation or degree of concentration around this average, while Cases 7–9 do. The MLE formulation of the problem can give more insight into the nature of the circle data by estimating the parameters in the distributions for Θ and ρ, especially when utilizing circular distributions (e.g. von Mises) for modeling the angular orientation.

A few of the cases in Table 1 have been mentioned in previous papers. Mardia and Holmes (1980) briefly comment on the Case 1 formulation of the circle fitting problem. Since they found this model inappropriate for the particular application considered in their paper, they did not develop or pursue this model. Berman and Culpin (1986) and Berman (1989) computed the asymptotic relative efficiency of the circle parameter estimates using least squares estimation versus maximum likelihood estimation under the Case 3 and 9 model assumptions. Their main findings were that the least squares estimates are inconsistent and inefficient when the radius has a large coefficient of variation and/or when the circle data is concentrated around a particular mean direction, i.e. not uniformly distributed around an entire circle. These two papers indicate when least squares estimation may not be adequate, but they do not provide any practical assistance for how the circle parameters should be estimated via maximum likelihood.

2.2 Maximum likelihood estimates

When the center (α, β) is known, and Θ and ρ are assumed to be independent, estimating the parameters in Ψ for the cases above is well understood. The observed (X, Y) coordinates, $(x_1, y_1), \ldots, (x_n, y_n)$, along with relationship (1) define a random sample of orientations and radii, $\theta_1, \ldots, \theta_n$ and $r_1, \ldots, r_n$, from which to estimate the model parameters for Θ and ρ respectively. Fisher (1993) and Mardia and Jupp (1999) cover maximum likelihood estimation of the von Mises parameters, and any number of mathematical statistics references cover maximum likelihood estimation of the uniform, normal, and gamma distributions' parameters. Table 2 summarizes the MLE's for the cases of distributions considered in this paper.

To elaborate on Table 2, the MLE's for the bounds on the uniform distribution for Θ are the minimum and maximum angles in the circular sense, in that they define the smallest arc containing all observed angles. The MLE for the mean direction μ_Θ of the von Mises distribution is the sample mean direction $\bar{\theta}$ of the observed angles. The estimate of the concentration parameter κ is $\hat{\kappa} = A^{-1}(\bar{R})$, where $A(\cdot)$ is the ratio

Table 2 Maximum likelihood estimates for parameters of the probability models in Table 1

Variable	Distribution	Parameter estimates
Θ	Uniform$(0, 2\pi)$	No parameters to estimate
	Uniform (l_Θ, u_Θ)	$\hat{l}_\Theta = \theta_{(1)}$ and $\hat{u}_\Theta = \theta_{(n)}$
	von Mises (μ_Θ, κ)	$\hat{\mu}_\Theta = \bar{\theta}, \hat{\kappa} = A^{-1}(\bar{R})$
ρ	Uniform (l_ρ, u_ρ)	$\hat{l}_\rho = r_{(1)}$ and $\hat{u}_\rho = r_{(n)}$
	Normal (μ_ρ, σ^2)	$\hat{\mu}_\rho = \bar{r}, \hat{\sigma}^2 = (n-1)s^2/n$
	Gamma (r, c)	No closed form solutions

of the first- and zeroth-order modified Bessel functions of the first kind, and $\bar{R}$ is the sample mean resultant length of the observed angles. In the case of a uniform distribution for ρ, the MLE's for the bounds are again the minimum and maximum values, the first and nth order statistics. Turning to the MLE's of the normal distribution's parameters, the sample mean $\bar{r}$ of the observed radii is the MLE for μ_ρ, and the sample variance s^2 is used to compute the estimate for σ^2. There are no closed form solutions for the maximum likelihood estimates of the gamma distribution's parameters; numerical methods must be employed (Choi and Wette 1969).

Maximum likelihood estimates are often biased, as is the case for all estimates in Table 2, other than $\hat{\mu}_\Theta$ and $\hat{\mu}_\rho$, the estimates of the means. Bias corrections can be made if desired. See Fisher (1993, p. 88) for bias correction of $\hat{\kappa}$ and Choi and Wette (1969) for a discussion of the bias of the MLE's for the gamma distribution's parameters. The unbiased estimate of the normal distribution's variance σ^2 is the sample variance s^2. Unbiased estimates of the upper and lower bounds of a uniform distribution can be obtained from the MLE's as follows. If in general $X_{(1)}$ and $X_{(n)}$ are the first and nth order statistics for a random sample of n random variates from a Uniform(l, u) distribution, then it can be shown that unbiased estimates for l and u can be obtained from the MLE's by taking:

$$\hat{l} = X_{(1)} - \left[\frac{X_{(n)} - X_{(1)}}{n-1}\right]$$
$$\hat{u} = X_{(n)} + \left[\frac{X_{(n)} - X_{(1)}}{n-1}\right]$$

As shown above, if the center of the circle around which the data is distributed is known, maximum likelihood circle fitting does not pose many challenges. However, in practice, the center may not be known and also needs to be estimated. Finding the MLE's of the center location and the parameters of the distributions in Table 1 is now much more complicated than just using the estimates in Table 2. Closed from solutions to finding the MLE's for α, β, and Ψ are not available for any of the 9 cases considered. Even Case 1, which at first glance may seem trivial, does not have a closed form solution when the center is unknown. For Case 1, it can be shown that finding the MLE's for α, β, l_ρ, and u_ρ is related to finding the minimum-width annulus containing the observed data. Some references discussing this non-trivial problem of computational geometry include Rivlin (1979), Agarwal et al. (2000) and Yu et al.

(2008). An additional problem arising from the complexity of the likelihood function is that typical numerical algorithms, such as Newton-Raphson, when applied to circle fitting are prone to complications such as converging to a local maximum and diverging to infinity (Zelniker and Clarkson 2006).

In Sect. 2.3 below, I will propose an intuitive Monte Carlo approach to finding the MLE's of the circle parameters. This approach has the appeal that it utilizes computational muscle to reduce the theoretical nature of the problem to being only as complicated as finding the MLE's for the parameters of the marginal distributions of Θ and ρ as given in Table 2.

2.3 Monte Carlo maximum likelihood circle fitting

If not for the unknown center of the circle (α,β), the maximum likelihood estimation of the parameters in Ψ is well-documented as described above (Table 2). However, with an unknown center the log-likelihood function, in terms of all parameters in Ψ and α and β, becomes very complicated. Instead of having to derive and use its gradient and Hessian matrix in an estimation algorithm that may not converge to a global maximum anyway, one can employ a Monte Carlo search technique that simplifies the estimation task considerably.

Robert and Casella (1999, p. 194) discuss "Stochastic Exploration" as a means of finding the solution to

$$\max_{\theta \in \Theta} h(\theta)$$

for a general parameter space Θ and objective function h (using their notation). Provided the parameter space is bounded, a sample from a uniform distribution over the parameter space can be obtained, $u_1, \ldots, u_m$. The maximum of $h(\theta)$ can be approximated by $h_m^* = \max(h(u_1), \ldots, h(u_m))$. This estimate converges as m goes to infinity.

For Monte Carlo circle fitting, it is not necessary to conduct stochastic exploration over the entire parameter space for Ψ, α and β. Instead, it suffices to simply take a large random sample of possible center locations. Conditional on a candidate center, one can compute the MLE's for the parameters in Ψ as given in Table 2. Approximate MLE's for Ψ, α, and β are then obtained by taking the combination of the candidate center location and its associated MLE's for Ψ that maximizes the log-likelihood function L.

Assuming one is sampling candidate centers from the correct parameter space for (α,β), as the number of sampled candidate centers goes to infinity, this method will produce parameter estimates that in the limit maximize the log-likelihood function. To locate a reasonable neighborhood from which to sample candidate center locations, one could start with a region centered around the least squares estimated center, or one could use an ocular estimate of the center location.

To be certain that the candidate centers are being sampled from the correct parameter space, a rather large neighborhood to sample from may be needed. This is especially the case when data is observed only for a small arc of a circle. It may not even be clear on which side of the data the center lies. One may need to sample an enormous number of candidate centers to obtain a suitably precise estimate of the MLE's.

Another strategy is to begin with a rather large area from which to initially sample candidate center locations, and then iteratively sample from smaller and smaller subregions; in a sense, zooming in on the center location that yields the largest overall log-likelihood. The theoretical convergence of the parameter estimates will be lost, as this algorithm can get trapped in a part of the parameter space for the circle center that does not contain the MLE's. However, in practical terms, the gain in precision may be worth the risk, especially if multiple implementations of this method lead to similar results. The following is an outline of such a Monte Carlo search algorithm for estimating the maximizers of the log-likelihood function L with respect to the unknown parameters Ψ, α, and β:

1. Obtain an initial estimate of the center location. Denote this by $(\hat{\alpha}_0, \hat{\beta}_0)$. Fixing α and β at these initial estimates, find the MLE's of the parameters in Ψ according to Table 2. Denote these by $\hat{\Psi}_0$. Let L_0 be the log-likelihood function evaluated for this set of parameter estimates.
2. Obtain a uniform random sample of k candidate centers in a circular neighborhood of radius δ around the initial estimate of the center $(\hat{\alpha}_0, \hat{\beta}_0)$.
3. For each of the k candidate centers, find the MLE's of the parameters in Ψ according to Table 2. Let L_1 be the maximum log-likelihood obtained by one of the k candidate centers and their associated MLE's of Ψ. Let $\hat{\Psi}_1, \hat{\alpha}_1, \hat{\beta}_1$ denote those maximum likelihood estimators.
4. Obtain a uniform random sample of k candidate centers in a circular neighborhood of radius $\delta/2$ around the current estimate of the center, $(\hat{\alpha}_1, \hat{\beta}_1)$. Proceed as in (3) to find updated estimates $\hat{\Psi}_2, \hat{\alpha}_2, \hat{\beta}_2$, with associated maximized log-likelihood L_2.
5. Continue iteratively, each time narrowing the neighborhood's radius, until the change in the maximized log-likelihood is less than some predefined $\epsilon > 0$, i.e. until $L_i - L_{i-1} < \epsilon$.

In order to prevent the maximized log-likelihood from decreasing from one iteration to the next, it is advisable to include the previous iteration's optimal center in the current iteration's collection of candidate centers. In this fashion, the above algorithm will produce a sequence of estimators for which the maximized log-likelihood cannot decrease from one iteration to the next.

The rate at which the neighborhood radii decrease from one iteration to the next is an optional parameter of the algorithm. Decreasing the size of the neighborhood more slowly will reduce the chance of converging to a local instead of a global maximum - at the cost of requiring more computations to obtain the same precision.

An Akaike information criterion can be used to compare model fits between the 9 model assumption cases proposed in this paper (or others). $AIC = -2L + 2k$, where k is the number of parameters estimated in the model for the data and L is the maximized log-likelihood function. A smaller value for AIC among a set of candidate models suggests a relatively better fit. AIC is known to be negatively biased (Hurvich and Tsai 1989), especially so when the number of model parameters is large relative to the sample size. A bias-corrected AIC could be pursued, but is left for future work.

Some examples in Sect. 3 will illustrate the above procedure. First, I will propose how to test the key model assumptions, including the independence between radius and angle.

2.4 Testing model assumptions

One of the assumptions underlying the model presented in Sect. 2.3 is that of independence between the radii and angles of the observed coordinates. The radii and angles are not directly observable, as they are defined by the unknown center of the circle. So, it is not possible to test the assumption of independence with only the observed (X, Y) coordinate data. After having estimated the model parameters, however, it is possible to use the estimated center to compute estimated radii and angles with which to test the independence assumption.

Given the observed (X, Y) coordinates $(x_1, y_1), \ldots, (x_n, y_n)$, and the maximum likelihood estimate of the circle center $(\hat{\alpha}, \hat{\beta})$, estimated angles and radii can be computed using the relationship in (1).

$$\hat{\theta}_i = \tan^{-1}\left(\frac{y_i - \hat{\beta}}{x_i - \hat{\alpha}}\right) \quad \text{and} \quad \hat{r}_i = \sqrt{(x_i - \hat{\alpha})^2 + (y_i - \hat{\beta})^2}.$$

The estimated pairs $(\hat{\theta}_1, \hat{r}_1), \ldots, (\hat{\theta}_n, \hat{r}_n)$ can then be used to test the assumption of independence between Θ and ρ. A linear-circular correlation coefficient can be used to test for an association. One such statistic is the linear-circular rank correlation coefficient, first presented in Mardia (1976) and also described in Mardia and Jupp (1999, p. 246) and Fisher (1993, p. 140). In the examples below, I will use the D_n version of the correlation coefficient, which is scaled to be in [0, 1], with values close to 0 indicating little association and values near 1 indicating strong association.

The model assumptions that the radii and angles vary according to the probability distributions given by a particular case in Table 1 can be verified once a model has been fit, using the estimated pairs of radii and angles as well. Standard goodness of fit tests, e.g. Kolmogorov-Smirnov, χ^2, and probability plots can be used to assess the adequacy of the normal, uniform, and gamma distributions. Mardia and Jupp (1999, Chap. 6) discuss goodness of fit tests for the circular uniform and von Mises distributions.

3 Examples

First, I will re-examine two archaeological examples that have been used in previous circle fitting papers: the ancient stone rings at Brogar, Scotland (Anderson 1981; Berman 1983) and archaeological excavations identifying an ancient Greek racecourse (Rorres and Romano 1997; Schaffrin and Snow 2010). Thereafter, I will illustrate the utility of some of the other distributional cases using simulated data.

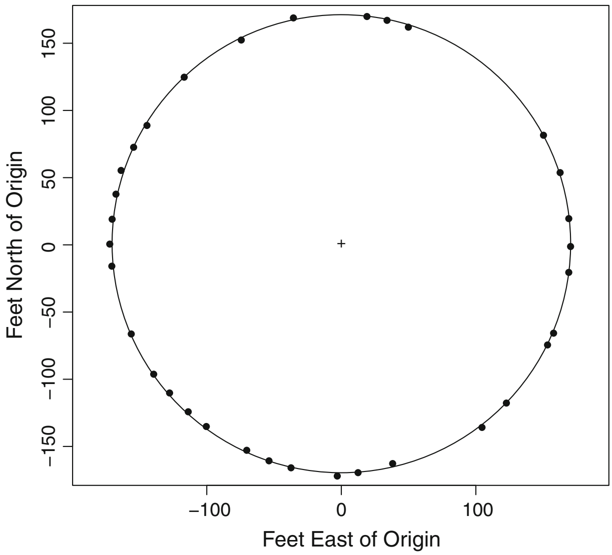

Fig. 2 Brogar stone ring with outliers removed ($n = 33$); Monte Carlo MLE circle, Case 2 fit

All computations were done with R version 2.11.1 (R Development Core Team 2010). The user CPU times that are provided are based on my experience on a Dell Optiplex 780, with an Intel® Core™ 2 Duo Processor E8400 with 3 GHz and 1.94 GB of RAM. I would expect significantly better performance with optimized code called from C++.

3.1 Stone ring at Brogar

The stone ring at Brogar consists of 35 stones in a circular formation (Thom and Thom 1973). There are two stones that have been considered outliers by some authors. I will also remove them from my analysis. Figure 2 displays the data and an MLE circle fit using the Monte Carlo method outlined in Sect. 2.3. One aspect of this stone circle that has been of interest in past literature is its diameter, which is believed to be related to the so-called megalithic yard.

Anderson (1981) used a maximum likelihood approach for analyzing this data as well. Her model assumptions and estimation approach is fundamentally different than what is presented in this paper. However, in the case of her Brogar ring data analysis, her model assumptions are similar to my Case 2 assumptions, allowing for comparisons to be drawn.

The MLE circle I fit to the Brogar ring data was obtained as follows. Case 2 model assumptions on ρ and Θ were made so that my results could be compared to Anderson (1981). To get an initial estimate for where the circle center is likely to be, a least squares fit was obtained, yielding the estimated center of $(\hat{\alpha}, \hat{\beta}) = (0.036, 0.844)$. The least squares estimated radius was $\hat{\rho} = 170.408$. Next, $k = 10,000$ circle centers were randomly generated from a neighborhood of radius $\delta = 2$ feet around the least

Table 3 Estimated circle parameters for Brogar stone ring data: Case 2 Monte Carlo MLE fit; ML estimates of Anderson (1981); least squares estimates

Case	Parameter estimates		
	Center (α, β)	Radius (μ_ρ, σ)	AIC
2	(0.035, 0.842)	(170.408, 1.161)	571.91
Anderson	NA	(170.41, 1.1651)	NA
LS	(0.036, 0.844)	Radius = 170.408	NA

squares circle center. Then, at each subsequent iteration the neighborhoods' radii were reduced by half. Once the maximized log-likelihood function changed by less than $\epsilon = 1 \times 10^{-10}$, the algorithm was assumed to have converged, resulting in the MLE's found in Table 3.

My Case 2 MLE's of the circle center are very close to the least squares estimates, and my estimated mean direction is virtually identical to the least squares estimated radius. The results presented in Anderson (1981) are very similar to the Monte Carlo MLE results I obtained. She did not report estimated centers, presumably because her focus was on estimating the diameter.

Anderson (1981) presented a 95% confidence interval for the diameter of the circle: 340.82 ± 0.56 feet. Her margin of error was computed using approximations of the asymptotic properties of the MLE's. I used a simple non-parametric, percentile bootstrap to find the confidence interval for the diameter, resampling 5,000 times: (340.028, 341.617). My bootstrap confidence interval is wider than Anderson (1981), 1.589 versus 1.12 feet, indicating that her results using large-sample asymptotics and approximations may be underestimating the sampling variability of the estimates just a bit.

Additional bootstrap methods could be used to make inference about the various circle parameters. Though not of central interest in the Brogar stone circle study, in general, it may be useful to construct 95% confidence regions for the circle center.

The assumption of independence between Θ and ρ needed for my maximum likelihood circle fit was assessed using the correlation coefficient suggested in Sect. 2.4. For the Brogar ring data, the correlation coefficient between the estimated radii and angles was $D_n = 0.0009$, with an approximate p value of 0.986 testing the null hypothesis of no association. There is no reason to believe the independence assumption is invalid based on this statistic.

The Case 2 Monte Carlo MLE procedure applied to the Brogar data required 4 iterations to converge. With 10,000 random candidate circle centers used at each iteration, in total 40,000 candidate circle centers were randomly drawn. The CPU running time to arrive at the MLE estimates was 9.70 s.

In order to see how robust the Monte Carlo MLE algorithm is for this data set, I repeated the estimation process 1,000 times (Table 4). The 95% mid-range for each quantity in Table 4 is the length of the interval that contains the middle 95% of all values obtained over the 1,000 simulations. For example, among the 1,000 estimates obtained for the mean of ρ, the middle 95% of the estimates were all within 3.083×10^{-5} of one another. There is very little variation in the parameter estimates obtained. This may

Table 4 Robustness of Brogar stone ring fitted circle: summary of 1,000 Case 2 Monte Carlo MLE fits

Parameter	ML estimates				
	Min	Max	95% mid-range	Mean	SD
α	0.0341	0.0350	0.000179	0.0346	0.0000507
β	0.842	0.843	0.000177	0.842	0.0000582
μ_ρ	170.408	170.408	0.0000308	170.408	0.00000753
σ	1.161	1.161	0.000000211	1.161	0.0000000521
AIC	571.911	571.911	0.000000709	571.911	0.000000380
Iterations	2	11	8	6.8	2.30

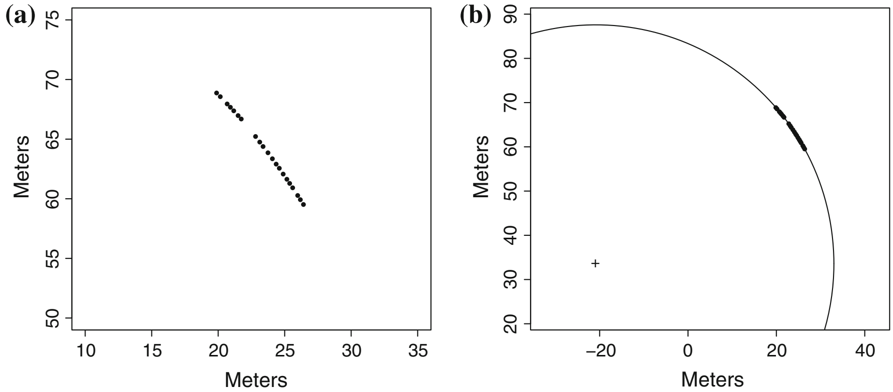

Fig. 3 **a** Starting line for an ancient Greek stadium in Corinth ($n = 21$); **b** Monte Carlo MLE circle, Case 5 fit

not be surprising, since a good starting value estimate of the center (the least squares circle center) was in hand, and the circle seems very well-defined. Later examples will investigate the robustness of the algorithm in less ideal conditions.

3.2 Ancient Greek stadium in Corinth

Next, I will consider the archeological data denoting the starting line of a circular track in an ancient Greek stadium (Rorres and Romano 1997; Schaffrin and Snow 2010). This data, along with my eventual Monte Carlo MLE fitted circle, can be seen in Fig. 3.

Because the data only covers a very small arc of an entire circle, some estimation algorithms have difficulty converging, or may be highly dependent on initial estimates of the circle parameters. Anderson (1981) points out this complication for maximum likelihood estimators when data lies only in a small arc. Berman (1989) shows that least squares circles are inconsistent and inefficient in this setting. Using least squares estimation on the Corinth data set I found vastly differing estimates, depending on the starting values I used for the parameter estimates. I also obtained different estimates depending on which type of optimization algorithm I requested the software to use.

Table 5 Estimated circle parameters for Corinth data: Cases 2 and 5 Monte Carlo MLE fit; least squares estimates (Rorres and Romano 1997)

Case	Parameter estimates			AIC
	Center (α, β)	Radius (μ_ρ, σ)	Angle (l_θ, u_θ)	
2	$(-20.190, 34.140)$	$(53.048, 0.013)$	NA	129.55
5	$(-20.943, 33.616)$	$(53.963, 0.013)$	$(0.490, 0.723)$	-8.47
LS	$(-20.940, 33.618)$	Radius = 53.960	NA	NA

My Monte Carlo search algorithm encountered no such difficulties, provided I used a large enough δ neighborhood from which to randomly sample candidate circle centers.

Using the least squares estimated center as a reference, the initial estimate for the center was taken to be $(-19, 35)$. Taking $k = 10,000$, $\delta = 5$, and $\epsilon = 1 \times 10^{-10}$, I obtained Monte Carlo MLE parameter estimates for Cases 2 and 5 (Table 5). In estimating the lower and upper bounds on the distribution of the orientation Θ for Case 5, I have reported unbiased estimators derived from the MLES's, and not just the MLE's themselves. Given that the data does not span an entire circle, it makes sense that Case 5 has a far better fit than Case 2 in terms of AIC.

The Case 5 MLE circle center was used to compute the estimated radii and angles of the data, from which a correlation coefficient of $D_n = 0.184$ with p value 0.163 was computed. There is insufficient evidence to suggest a problem with the independence assumption between Θ and ρ underlying the Case 5 model fit.

The Case 5 Monte Carlo MLE parameter estimates of α, β, and μ are identical up to two decimals to the least squares estimates reported by others (Rorres and Romano 1997; Schaffrin and Snow 2010). The fact that the data is limited to a very small arc of the circle is not reflected in the least squares estimators, while it is accounted for in the Case 5 MLE fit. This example illustrates that the MLE approach can sometimes provide more insight into the nature of the data than a least squares circle can.

The CPU running time for the Case 5 Monte Carlo fit was 24.58 s, requiring 4 iterations to converge. In total, again, 40,000 candidate circle centers were sampled in the course of the algorithm.

To check robustness, a total of 1,000 Case 5 MLE circles were fit using the Monte Carlo MLE algorithm. For this simulation I did not use the least squares circle center as a starting value. Instead, I pretended I had limited information about where the optimal center might be. I used an initial center of $(-200, -100)$, and began the algorithm by sampling candidate centers from a neighborhood of radius $\delta = 500$ around $(-200, -100)$. Recall that originally I used a starting center estimate of $(-19, 35)$ with $\delta = 5$. The simulation shows that the larger sampling neighborhood had an impact on the robustness of the circle parameter estimates. In particular, the center and mean radius estimates appear to have non-trivial variability. The other parameters are much more stable (Table 6).

The simulation was rerun using $(-19, 35)$ as the initial estimate for the circle center, and the region with radius 5 around this point as the starting sample space for candidate centers, as was done in the original model estimation for this data set.

Table 6 Robustness of Corinth data fitted circle: summary of 1,000 Case 5 Monte Carlo MLE fits

Parameter	ML estimates				
	Min	Max	95% mid-range	Mean	SD
α	−21.709	−19.6574	0.208	−20.938	0.0799
β	33.082	34.503	0.146	33.620	0.0557
μ_ρ	52.405	54.895	0.255	53.957	0.0972
σ	0.0130	0.0133	0.0000122	0.0130	0.0000124
l_θ	0.498	0.502	0.000505	0.501	0.000197
u_θ	0.710	0.716	0.000460	0.712	0.000189
AIC	−8.472	−7.408	0.0393	−8.467	0.0396
Iterations	3	22	16	11.9	4.98

The results (not tabulated) are much more stable as one might expect. The 95% mid-ranges and standard deviations of the parameter estimates all decrease by about two orders of magnitude.

This example demonstrates that more robust MLE circle estimates can be obtained by using the least squares circle or other prior information as a guide. In practice, if little prior information is at hand, a simulation as resulted in Table 6, could be done as an initial analysis to get a sense of where the circle center lies. For example, I could have followed up the first simulation with another Monte Carlo MLE circle fit using a smaller δ around the average center coordinates $(-20.938, 33.620)$ as my starting center estimate.

3.3 Simulated data examples

3.3.1 Case 4 data

The first simulated data set I will present consists of 50 (X, Y) coordinates, randomly generated using a center at (10, 10) and the model assumptions of Case 4: specifically, observations have radii that are uniformly distributed on (6, 8), and orientations that are uniformly distributed on $(\pi/2, \pi)$. The data along with the least squares and a Monte Carlo MLE fitted circle can be seen in Fig. 4.

Cases 4–9 would be more or less reasonable candidate models to try, if one were interested in finding the maximum likelihood circle for this data. Cases 1–3 are ill-advised, given the data clearly lies in an arc of only about 45°. I used the Monte Carlo MLE algorithm with the following specifications: $k = 10,000$, $\delta = 10$, and $\epsilon = 1 \times 10^{-10}$, and the initial center was again selected to be near the least squares estimated center, at (9, 11). The least squares and Case 4 Monte Carlo MLE estimates are given in Table 7. The Case 4 estimates for the uniform distribution parameters have been bias corrected. From the estimated limits of the uniform distribution for the radius, the average radius is estimated to be 7.188. Contrasting this estimate with the least squares estimated radius of 5.879, one can see numerically what Fig. 4a indicates:

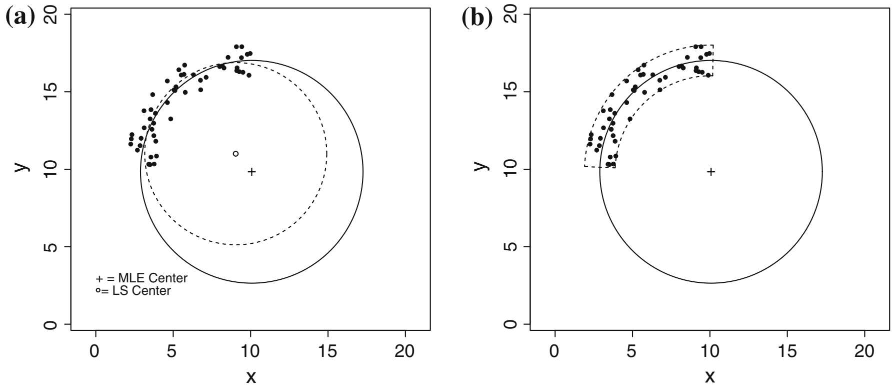

Fig. 4 Case 4 simulated data ($n = 50$); **a** Monte Carlo MLE fit, Case 4 (*solid line*), and LS circle (*dashed line*); **b** Monte Carlo MLE fit, Case 4, with estimated bounds on ρ and Θ

there is a substantial difference between the least squares and Monte Carlo MLE circles in this example.

Not only are the Case 4 MLE parameters much closer to the parameter values used to generate the random data, but the MLE approach provides more insight into the nature of the data than the least squares circle could provide; or for that matter, what the maximum likelihood models presented in prior literature could provide. It may be quite useful in certain applications to be able to estimate the bounds of both ρ and Θ, as I have done here. The bounds on the distributions of ρ and Θ can be illustrated graphically as in Fig. 4b.

I conducted a simulation study to see how the Monte Carlo MLE algorithm performs on average, and to assess how the parameter estimates and model fits compare across the 9 cases, when the true model parameters are known. One thousand data sets with 50 observations and the same Case 4 model assumptions as described above were randomly generated. The Case 4 entry in Table 8 shows the results of the simulation when the same model is fit to the data as was used to generate the data. The Case 4 fits had the lowest AIC on average, indicating that AIC can be a useful measure to discriminate between the model fits of various cases. This simulation also shows that circle fitting methods that assume the data is distributed around the entire circle may estimate the circle center rather poorly. The estimated centers for Cases 4–9 are all very similar and close to the theoretical value. It is these cases that do not assume the data is distributed around the entire circle; Cases 1–3 do make this assumption. I expect that least squares circles would also have estimated the circle centers less accurately on average, as was already indicated in Table 7.

The only convergence problem that I had with the Monte Carlo MLE algorithm (in this example, as well as all others) was when using cases involving the gamma distribution (Cases 3, 6, 9). As mentioned above, estimating the gamma distribution's parameters is an iterative process itself. A poor candidate center estimate will sometimes prevent this iterative estimation process from converging. I took a lack of convergence of the estimation algorithm for the gamma distribution's parameters as evidence that the candidate center that led to this non-convergence could not possibly

Table 7 Estimated circle parameters for Case 4 simulated data: Case 4 Monte Carlo MLE fit; least squares estimates

Case	Parameter estimates			AIC
	Center (α, β)	Radius (l_ρ, u_ρ)	Angle (l_θ, u_θ)	
4	(10.074, 9.835)	(6.198, 8.177)	(1.554, 3.101)	313.21
LS	(9.039, 11.002)	Radius = 5.879		

Table 8 Mean (SD) of estimated circle parameters for 1,000 data sets simulated under Case 4 assumptions

Case	Center		Case specific model parameters	AIC	Iterations
	α	β			
1	7.79 (1.09)	12.21 (1.09)	$l_\rho = 3.03$ (1.47), $u_\rho = 5.89$ (0.78)	432.98 (8.69)	16.90 (11.10)
2	7.35 (0.57)	12.66 (0.58)	$\mu_\rho = 3.95$ (0.58), $\sigma_\rho = 0.78$ (0.16)	440.12 (9.01)	11.76 (9.42)
3	7.33 (0.60)	12.69 (0.63)	$r = 30.10$ (16.50), $c = 0.19$ (0.13)	439.95 (9.23)	12.04 (9.75)
4	10.04 (0.51)	9.97 (0.50)	$l_\rho = 6.08$ (0.65), $u_\rho = 8.03$ (0.65)	309.84 (4.74)	14.68 (7.43)
			$l_\theta = 1.57$ (0.08), $u_\theta = 3.14$ (0.08)		
5	10.16 (0.79)	9.85 (0.78)	$\mu_\rho = 7.21$ (1.00), $\sigma = 0.56$ (0.04)	330.60 (7.50)	9.26 (5.48)
			$l_\theta = 1.58$ (0.11), $u_\theta = 3.13$ (0.11)		
6	10.17 (0.79)	9.84 (0.78)	$r = 173.27$, (54.64), $c = 0.04$ (0.0085)	330.63 (7.55)	9.40 (5.73)
			$l_\theta = 1.58$ (0.11), $u_\theta = 3.13$ (0.11)		
7	10.17 (0.54)	9.84 (0.54)	$l_\rho = 6.26$ (0.70), $u_\rho = 8.20$ (0.70)	331.05 (7.93)	14.76 (7.45)
			$\mu_\theta = 2.36$ (0.07), $\kappa = 5.54$ (1.30)		
8	10.47 (0.78)	9.54 (0.78)	$\mu_\rho = 7.62$ (1.00), $\sigma = 0.56$ (0.04)	351.65 (9.87)	10.15 (7.05)
			$\mu_\theta = 2.36$ (0.07), $\kappa = 6.18$ (1.79)		
9	10.48 (0.78)	9.53 (0.78)	$r = 191.95$ (55.80), $c = 0.04$ (0.0076)	351.66 (9.92)	10.15 (7.12)
			$\mu_\theta = 2.36$ (0.07), $\kappa = 6.20$ (1.79)		

be optimal. Accordingly, the log-likelihood for such a candidate center was defined to be negative infinity, so that such a candidate center would not be selected as optimal in the overall Monte Carlo MLE circle fitting process.

3.3.2 Case 1 data

My final example illustrates the relationship between the Case 1 MLE circle fit and the minimum-width annulus encompassing a set of points in $\Re^2$. For Case 1 model assumptions, the likelihood function of the observed (X, Y) coordinates is

$$lik(\alpha, \beta, l_\rho, u_\rho | \mathbf{x}, \mathbf{y}) = \left[\frac{1}{2\pi(u_\rho - l_\rho)}\right]^n \prod_{i=1}^{n} \frac{1}{\sqrt{(x_i-\alpha)^2 + (y_i-\beta)^2}}, \tag{3}$$
$$\text{where } l_\rho \le \sqrt{(x_i-\alpha)^2 + (y_i-\beta)^2} \le u_\rho \text{ for all } i = 1, \ldots n.$$

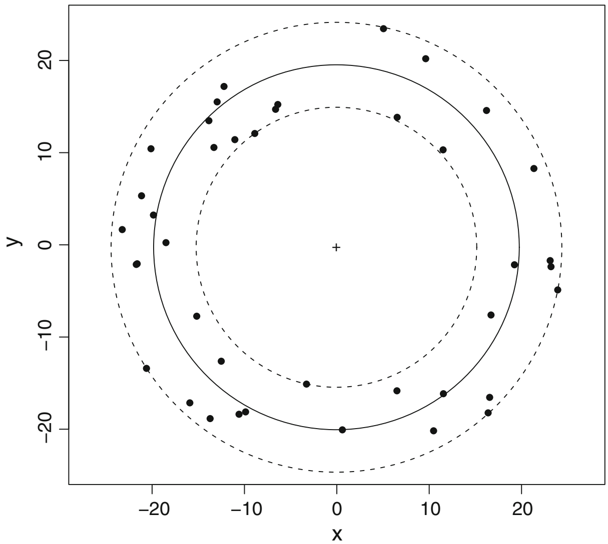

Fig. 5 Case 1 simulated data ($n = 40$); Monte Carlo MLE fit, Case 1, yielding an annulus containing the data

If the center (α, β) is known, maximizing this likelihood function is equivalent to finding values of l_ρ and u_ρ that are as close to one another as possible, while still having all points, i.e. their radii $r_i = \sqrt{(x_i - \alpha)^2 + (y_i - \beta)^2}$, within the annulus defined by l_ρ and u_ρ. This is equivalent to finding the minimum-width annulus for a set of points given a known center. This is accomplished by using an inner radius (i.e. $\hat{l}_\rho$) equal to the distance from the center to the nearest point to the center, and an outer radius (i.e. $\hat{u}_\rho$) equal to the distance from the center to the farthest point from the center.

If the center is not known, the maximum likelihood solution is more complicated than just finding the minimum-width annulus for the set of points. Maximizing (3) requires all points to be within an annulus, but defining the annulus requires consideration of the product of the nth power of the width of the annulus and the data points' radii. This is even more complicated than just finding a minimum width annulus for data in $\Re^2$, which has no closed form solution itself, and requires computational geometry algorithms and software to find approximate solutions (Rivlin 1979; Agarwal et al. 2000; Yu et al. 2008; CGAL 2011). However, the Monte Carlo MLE method to maximize the likelihood function for Case 1 in (3) is much more tractable and easily implemented. A large random sample of candidate centers is generated. For each of these centers the minimum-width annulus for the data is obtained. The likelihood function (3) can now be evaluated for all of these candidate sets of $\alpha, \beta, l_\rho, u_\rho$. The set that maximizes (3) is used as the estimated MLE's.

For example, 40 observations were simulated under Case 1 model assumptions, with a circle center at $(0, 0)$, radii uniformly distributed between 15 and 25, and orientations uniformly distributed around the entire circle (Fig. 5). The

Table 9 Estimated circle parameters for Case 1 simulated data

Case	Parameter estimates			AIC
	Center (α, β)	Radius (l_ρ, u_ρ)	Radius (μ_ρ, σ)	
1	$(-0.040, -0.262)$	(15.188, 24.400)	NA	572.67
2	$(0.470, -1.039)$	NA	(20.378, 2.843)	592.49
LS	$(0.586, -1.041)$	Radius = 20.395		

least squares circle has an estimated radius of 20.395, and an estimated center at $(0.586, -1.041)$. A Case 1 Monte Carlo MLE circle was fit using an initial estimated center near the least squares estimate: $(0.50, -1.00)$. The choice for the other inputs for the Monte Carlo MLE algorithm were $\delta = 10$, $k = 10,000$, and $\epsilon = 1 \times 10^{-10}$. A bias correction was not done on the uniform distribution parameters for the radius, in order for these parameter estimates to define an annulus for which the nearest/farthest points from center lie on the inner/outer circles of the annulus.

Averaging the uniform distribution's estimated bounds, $\hat{l}_\rho$ and $\hat{u}_\rho$, defines the MLE circle with an estimated radius of 19.794 (Table 9). A circle with this radius is shown in Fig. 5, along with dashed lines defined by the MLE's of the lower/upper bounds of the uniform distribution for the radius. For comparison, a Case 2 model fit was also obtained, yielding a worse fit according to AIC.

4 Conclusion

In many applications of circle fitting a maximum likelihood formulation of the problem can provide a deeper understanding of the underlying distributions that gave rise to a circularly patterned data set than a least squares approach can. Conventional approaches to the maximum likelihood solutions have been theoretically complex and have been shown to be computationally unreliable in some cases. In this paper I have presented a Monte Carlo algorithm that is intuitive and relatively straightforward in the theoretical statistics sense. I believe practitioners of statistics will find this approach appealing in its simplicity. The Monte Carlo MLE algorithm can be augmented with resampling techniques, such as the bootstrap, for additional inference.

I have written several functions in R to generate the circular plots and to estimate the least squares and maximum likelihood circles. These functions will be made available in an R package at CRAN. Until then, individuals are welcome to contact me if they are interested in the R code.

Acknowledgments Many thanks to an editor and associate editor for helpful suggestions that greatly improved this manuscript.

References

Agarwal PK, Aronov B, Har-Peled S, Sharir M (2000) Approximation algorithms for minimum-width annuli and shells. Discrete Comput Geom 24:687–705

Ahn SJ, Jeong WB, Yoo WS (2004) Unbiased expression of FRF with exponential window function in impact hammer testing. J Sound Vib 277:931–941
Anderson DA (1981) The circular structural model. J Royal Stat Soc Ser B 43(2):131–141
Berman M (1983) Estimating the parameters of a circle when angular differences are known. J Royal Stat Soc Ser C 32(1):1–6
Berman M (1989) Large sample bias in the least squares estimators of a circular arc center and its radius. Comput Vis Graph Image Process 45:126–128
Berman M, Culpin D (1986) The statistical behaviour of some least squares estimators of the centre and radius of a circle. J Royal Stat Soc Ser B 48(2):183–196
CGAL (2011) Computational Geometry Algorithms Library. http://www.cgal.org. Accessed 24 August 2011
Chan NN (1965) On circular functional relationships. J Royal Stat Soc Ser B 27(1):45–56
Chan YT, Thomas SM (1995) Cramer-Rao lower bounds for estimation of a circular arc and its radius. Graph Models Image Process 57(6):527–532
Chan YT, Lee BH, Thomas SM (2005) Approximate maximum likelihood estimation of circle parameters. J Optim Theory Appl 125(3):723–734
Chernov N, Lesort C (2005) Least squares fitting of circles. J Math Imaging Vis 23(3):239–252
Chernov N, Sapirstein PN (2008) Fitting circles to data with correlated noise. Comput Stat Data Anal 52(12):5328–5337
Choi SC, Wette R (1969) Maximum likelihood estimation of the parameters of the gamma distribution and their bias. Technometrics 11(4):683–690
Coope ID (1993) Circle fitting by linear and nonlinear least squares. J Optim Theory Appl 76(2):381–388
Ertürk S (2006) Nonintrusive iris image extraction for iris recognition-based biometric identification. Circuits Syst Signal Process 25(3):405–419
Fisher NI (1993) Statistical analysis of circular data. University Press, Cambridge
Hurvich CM, Tsai CL (1989) Regression and time series model selection in small samples. Biometrika 76(2):297–307
Janóczki M, Jakab László (2010) X-ray solder alloy volume measurement (XSVM) in pin-in-paste technology (PIP). Solder Surf Mount Technol 22(1):26–40
Kasa I (1976) A circle fitting procedure and its error analysis. IEEE Trans Instrum Meas 25:8–14
Mardia KV (1976) Linear-circular correlation coefficients and rhythmometry. Biometrika 63:403–405
Mardia KV, Holmes D (1980) A statistical analysis of megalithic data under elliptic pattern. J Royal Stat Soc Ser A 143(3):293–302
Mardia KV, Jupp PE (1999) Directional statistics. Wiley, Chichester
R Development Core Team (2010) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org
Rivlin TJ (1979) Approximation by circles. Computing 21(2):93–104
Robert CP, Casella G (1999) Monte Carlo statistical methods. Springer, New York
Rorres C, Romano DG (1997) Finding the center of a circular starting line in an ancient Greek stadium. SIAM Rev 39(4):745–754
Schaffrin B, Snow K (2010) Total least-squares regularization of Tykhonov type and an ancient ractrack in Corinth. Linear Algebra Appl 432(8):2061–2076
Thom A, Thom AS (1973) A megalithic lunar observatory in Orkney: the Ring of Brogar and its cairns. J Hist Astron 4(2):111–123
Thomas L, Shaffer CA, Mili L, Thomas E (2007) Automated detection of severe surface defec ts on barked hardwood logs. For Prod J 57(4):50–56
Umbach D, Jones K (2003) A few methods for fitting circles to data. IEEE Trans Instrum Meas 52(6):1881–1885
Yu H, Agarwal PK, Poreddy R, Varadarajan KR (2008) Practical methods for shape fitting and kinetic data structures using coresets. Algorithmica 52(3):378–402
Zelniker EE, Clarkson IVL (2006) Maximum-likelihood estimation of circle parameters via convolution. IEEE Trans Image Process 15(4):865–876