A STATISTICAL ANALYSIS OF LEAST-SQUARES CIRCLE-CENTRE ESTIMATION

Emanuel E. Zelniker and I. Vaughan L. Clarkson

Intelligent Real-Time Imaging and Sensing Group
School of Information Technology & Electrical Engineering
The University of Queensland
Queensland, 4072, AUSTRALIA
{zelniker,v.clarkson}@itee.uq.edu.au

ABSTRACT

In this paper, we examine the problem of fitting a circle to a set of noisy measurements of points on the circle’s circumference. An estimator based on standard least-squares techniques has been proposed by DELOGNE which has been shown by KÅSA to be convenient for its ease of analysis and computation. Using CHAN’s circular functional model to describe the distribution of points, we perform a statistical analysis of the estimate of the circle’s centre, assuming independent, identically distributed Gaussian measurement errors. We examine the existence of the mean and variance of the estimator for fixed sample sizes. We find that the mean exists when the number of sample points is greater than 2 and the variance exists when this number is greater than 3. We also derive approximations for the mean and variance for fixed sample sizes when the noise variance is small. We find that the bias approaches zero as the noise variance diminishes and that the variance approaches the CRAMÉR-RAO lower bound. We also show this through Monte-Carlo simulations.

1. INTRODUCTION

The accurate fitting of a circle to noisy measurements of points on its circumference is an important and much-studied problem in statistics. It has applications in many areas of research including archaeology [1], geodesy [2], microwave engineering [3] and computer vision and metrology [4].

The problem of obtaining an accurate circular fit, by which we mean the estimation of a circle’s centre and its radius, appears to have been first studied by THOM [1] in connection with measurements of ancient stone circles in Britain. He proposes an approximate method of least squares solution. In addressing a problem of ‘statistical geography’, ROBINSON [2] gives a complete formulation of the solution to the problem by the method of least squares.

The first detailed statistical analysis to be published appears to be that of CHAN [5]. He proposes a ‘circular functional relationship’, which we also use as the basis for our investigations. In this model, it is assumed that the measurement errors are instances of independent and identically distributed (i.i.d.) random variables. Additionally, the points are assumed to lie at fixed but unknown angles around the circumference, i.e., not only are the centre and radius of the circle unknown parameters to be estimated, but so are the angles of each circumferential point. He derives a method to find the maximum likelihood estimator (MLE) when the errors have a Gaussian distribution. This method is identical to the least-squares method of [2]. He also examines the consistency of the estimator.

A disadvantage of the MLE is that it is difficult to analyse. From a numerical point of view, another disadvantage is that the only known algorithms for computing the MLE are iterative. Furthermore, it has been reported that there are instances in which there is no minimum, but rather a stationary point, or several local minima in the likelihood function [6]. The difficulties with the MLE were recognised by KÅSA [3], who proposes using a simple estimator due to DELOGNE [7] which is relatively easy to analyse and also to compute and is based on least-squares techniques. This estimator has subsequently been independently rediscovered at least twice [8, 9].

BERMAN & CULPIN [6] have carried out a statistical analysis of the MLE and the DELOGNE-KÅSA estimator (DKE). Specifically, they prove some results regarding the asymptotic consistency and variance of the estimates. CHAN & THOMAS [10] have investigated the CRAMÉR-RAO lower bound (CRLB) for estimation in the circular functional model, but see also [11].

In this paper, we are interested in the properties of the DKE for fixed (small) sample sizes rather than its asymptotic properties. KÅSA himself carries out a ‘first-order’ or ‘small-error’ analysis of the estimator [3]. However, when the random variables that give rise to the errors are Gaussian, it can no longer be guaranteed that the errors will always be small, no matter how small the variance, and so the analysis becomes invalid. At the outset, it is not even clear whether the mean or variance of the estimator exists.

KÅSA states, by way of justifying the first-order analysis, that ‘it may be appreciated that [the expressions for the mean and variance of the estimator] are, in general, very hard to evaluate’. Nevertheless, in this paper, we demonstrate that, under certain conditions, the defining integrals are not wholly intractable. From analysis of the integrals, we set out conditions for which the mean and variance of the DKE for circle centre exist for fixed sample sizes under CHAN’s circular functional model with Gaussian errors. Where the mean exists, we show that the estimator is unbiased in the limit as noise variance approaches zero. Where the variance exists, we show that the variance approaches the CRLB as the variance approaches zero. We additionally demonstrate these theoretical results through Monte-Carlo simulations.

Emanuel Zelniker is additionally supported by a scholarship from the Commonwealth Scientific & Industrial Research Organisation.
2. BACKGROUND

2.1. Chan’s Circular Functional Model

We now briefly present Chan’s circular functional model [5]. In this model, we assume that the positions of $N$ points on the circumference of a circle are measured. The measurement process introduces random errors so that the Cartesian coordinates $(x_i, y_i)$, $i = 1, \ldots, N$ can be expressed as

$$x_i = a + r \cos \theta_i + \xi_i, \quad y_i = b + r \sin \theta_i + \eta_i.$$

Here, $(a, b)$ is the centre of the circle, $r$ is its radius, the $\theta_i$ are the angles around the circumference on which the points lie and the $\xi_i$ and $\eta_i$ are zero-mean and i.i.d. instances of random variables representing the measurement error. They are assumed to be normal with variance $\sigma^2$. In this paper, we explicitly exclude the possibility that $r = 0$ or $\theta_1 = \theta_2 = \ldots = \theta_N$.

Figure 1 shows some data with $N$ points for the circumference of a circular arc, $(x_1, y_1), \ldots, (x_N, y_N)$, displaced from the circumference by noise.

2.2. Maximum Likelihood Estimation

If we define $r_i(a, b) = \sqrt{(x_i - a)^2 + (y_i - b)^2}$, then Chan [5] showed that the MLE is

$$\hat{a}_{\text{ML}}, \hat{b}_{\text{ML}}, \hat{r}_{\text{ML}} = \text{arg min}_{(a, b, r)} \sum_{i=1}^{N} r_i(a, b) - r^2.$$

The difficulties with the MLE are that it is hard to analyse and also to compute numerically. Analytically, it is not certain that a global minimum exists, or whether there might be local minima [6]. Numerically, the only methods available for solution are iterative. This raises the usual issues with convergence and sensitivity to the initial solution estimate.

2.3. The Delogne-Kása Estimator

The analytical and numerical difficulties with the MLE in Section 2.2 led Kása [3] to propose the use of a modified estimator, originally due to Delogne [7] which we can write as

$$\hat{a}_{\text{DK}}, \hat{b}_{\text{DK}}, \hat{r}_{\text{DK}} = \text{arg min}_{(a, b, r)} \sum_{i=1}^{N} r_i^2(a, b) - r^2.$$

The linearisation which results from this formulation simplifies the analysis and the computation considerably. It can be shown that this estimator is a standard linear least-squares estimator. In terms of matrix algebra, we have

$$Z = (\hat{a}_{\text{DK}}, \hat{b}_{\text{DK}}, \hat{r}_{\text{DK}})^T = \frac{1}{2}(S + T)^\#(u + v).$$

(1)

Here, the superscript ‘$\#$’ represents the Moore-Penrose generalised inverse or pseudo-inverse and, for a matrix $A$ where $A^T A$ is non-singular, we may write $A^\# = (A^T A)^{-1} A^T$. Further,

$$S = (s_1, s_2) = PS', \quad T = PT', \quad u = Pu' \quad \text{and} \quad v = P \nu'$$

where

$$S' = \begin{pmatrix} a + r \cos \theta_1 & b + r \sin \theta_1 \\ \vdots & \vdots \\ a + r \cos \theta_N & b + r \sin \theta_N \end{pmatrix}, \quad T' = \begin{pmatrix} \xi_1 & \eta_1 \\ \vdots & \vdots \\ \xi_N & \eta_N \end{pmatrix}.$$

The elements of $u'$ and $v'$ are given by the expressions

$$u'_i = (a + r \cos \theta_i)^2 + (b + r \sin \theta_i)^2,$$

$$v'_i = 2(a + r \cos \theta_i) \xi_i + 2(b + r \sin \theta_i) \eta_i + \xi_i^2 + \eta_i^2$$

and $P$ is an $N \times N$ symmetric projection matrix defined so that $P = I - (11^T / N)$ where $I$ is the $N \times N$ identity matrix and $1$ is an $N$-dimensional column vector, all of whose entries are 1. Note that $\|P\|_2 = 1$.

3. ANALYSIS OF THE DELOGNE-KÁSA ESTIMATOR

We now turn our attention to the analysis of the DKE for fixed sample sizes. We are firstly interested in the question of whether the mean and variance exist. Later, we derive low-variance approximations for their values which are valid whenever they exist.

Before outlining the proofs for the main theorems in this section, we observe the following lemmas. The full proofs of all theorems and lemmas may be found in [12].

Lemma 1. The matrix $P$ has a singular-value decomposition of the form

$$P = YA_\Delta Y^T$$

where $Y$ is an orthogonal matrix and $\Delta = \text{diag}\{1, \ldots, 1, 0\}$.

Lemma 2. For any vectors $X, \mu \in \mathbb{R}^N$,

$$\exp \left(-\frac{\|X - \mu\|^2}{2\sigma^2}\right) \leq \exp \left(\frac{\|\mu\|^2}{2\sigma^2}\right) \exp \left(-\frac{\|X\|^2}{4\sigma^2}\right).$$

Corollary 1. If $X = (X_1, \ldots, X_N)^T$ is a multivariate normal random vector such that each $X_i \sim N(\mu_i, \sigma^2)$ is independent, then

$$E[\|f(X)\|^2] \leq 2N^2 \exp \left(-\frac{\|\mu\|^2}{2\sigma^2}\right) E[\|f(Y)\|^2],$$

where $\mu = (\mu_1, \ldots, \mu_N)^T$ and $Y = (Y_1, \ldots, Y_N)^T$ is a multivariate normal random vector such that each $Y_i \sim N(0, 2\sigma^2)$ is independent.
Definition 1. We say that an \( N \times n \) matrix \( X \) is a rectangular Gaussian matrix if each element is i.i.d. with identical variance \( \sigma^2 \) and \( E[X] = \mu \). We denote its distribution \( G(N, n, \mu, \sigma^2) \).

Theorem 1. The mean of the DKE for circle centre, as defined in (1), exists if the number of sample points on the circumference, \( N \), is greater than 2.

Outline of proof. If the variance \( \sigma^2 \) is zero then \( Z \) is deterministic and \( Z = \frac{1}{2} S^{\#} u \). In this case, the mean clearly exists, since the pseudo-inverse of \( S \) always exists. Hence, we restrict our attention to the case where \( \sigma^2 > 0 \).

In order to show that the expectation exists, it is necessary to show that \( E[||Z||^2] < \infty \). Let \( Y' \) be a random matrix with distribution \( G(N, 2, S', \sigma^2) \) such that \( Y' = S' + T' \). From the definition of expectation, \( E[||Z||^2] = \frac{1}{2} E[||(S + T')^\#(u + v)||^2] = \frac{1}{2} E[||(PY')^\#||^2] \), and we note that \( v \) is a function of \( Y' \), i.e., a function of the \( \xi \) and \( \eta \). Through the use of Corollary 1 and the sub-multiplicative inequality, we find that

\[
E[||Z||^2] \leq 2^{N-2} \exp \left( \frac{||S'||^2}{2\sigma^2} \right) E[||(PW')^\#||^2] \leq 2 \cdot \exp \left( \frac{||S'||^2}{2\sigma^2} \right) E[||(PW')^\#||^2] \tag{4}
\]

where \( || \cdot ||_F \) represents the Frobenius norm of its argument and \( W' \) is a random matrix like \( Y' \) but each element has zero mean and twice the variance, i.e., \( W' \sim G(N, 2, 0, 2\sigma^2) \).

Define \( \delta \) and \( \epsilon \) as the first \( N - 1 \) elements of \( Y'^T \xi \) and \( Y'^T \eta \), as defined in Lemma 1. Let us also define \( \xi \) and \( \eta \) to be the last (i.e., \( N^\# \)) elements of \( Y'^T \xi \) and \( Y'^T \eta \) so that

\[
\begin{bmatrix} \delta \\ \epsilon \\ \eta \end{bmatrix} = Y'^T \xi, \quad \begin{bmatrix} \xi \\ \eta \end{bmatrix} = Y'^T \eta \tag{5}
\]

Notice that \( Y'^T W' \sim G(N, 2, 0, 2\sigma^2) \). Then, by Lemma 1

\[
\Delta Y'^T W' = \begin{bmatrix} W \\ 0 \end{bmatrix}, \quad W = (\delta, \eta),
\]

and \( W \sim G(N - 1, 2, 0, 2\sigma^2) \). Also, 

\[
||(PW')^\#||^2 = ||W'||^2 \tag{6}
\]

We would now like to note that it is not very difficult to bound \( ||u + v(W')||^2 \) above by a degree two polynomial in \( ||W||^2, ||\xi|| \) and \( ||\eta|| \). Since these are independent random variables, we can take the expectation of \( ||\xi|| \) and \( ||\eta|| \) separately so that we can now rewrite (4) to show that

\[
E[||Z||^2] \leq 2^{N-2} \exp \left( \frac{||S'||^2}{2\sigma^2} \right) E[||(PW')^\#||^2] \tag{7}
\]

where \( p_1(||W||^2) \) is a degree two polynomial in \( ||W||^2 \) only. Consider the value of \( ||W||^2 \) and \( ||W'||^2 \), i.e.,

\[
||W||^2 = s_{\max}, \quad ||W'||^2 = \frac{1}{s_{\min}}.
\]

Now, \( s_{\max} \) and \( s_{\min} \) are the singular values of \( W \) and therefore are the square roots of the maximum and minimum eigenvalues of \( W^T W \). Then, \( W^T W \) has a Wishart distribution and from Muirhead [13, p. 106], the exact joint density function for the 2 singular values of \( W^T W \) can be written as

\[
P_{s_{\max},s_{\min}}(s_{\max}, s_{\min}) = \exp \left( -\frac{s_{\max}^2 + s_{\min}^2}{2} \right) \int_0^\infty \int_0^\infty p_1(s_{\max}) ds_{\max} ds_{\min}
\]

where \( K_{N,2} \) is a normalising constant. Looking at (7) and using (8), we can say that

\[
e[||W||^2 ||W'||^2] = \frac{1}{2\sigma^2} \int_0^\infty p_1(s_{\max}) ds_{\max} \int_0^\infty p_1(s_{\min}) ds_{\min}.
\]

It can now be seen from (11) that we have bounded \( E[||Z||^2] \) above a two-dimensional integral in \( s_{\max} \) and \( s_{\min} \). This integral is the product of a degree 2 polynomial of non-negative powers of \( s_{\max} \) and \( s_{\min} \) with an exponential of the negative square of \( s_{\max} \) and \( s_{\min} \) when \( N \geq 3 \). Such an integral is finite, e.g., see [14, §3.461].

For the remainder of this section, we omit the proofs of our theoretical results. Instead, in the next section, we will demonstrate them through simulation results. The full proofs of all theorems and lemmas may be found in [12], and all follow a similar theme to the proof of Theorem 1.

Theorem 2. The variance of the DKE for the circle centre, as defined in (1), exists if the number of sample points on the circumference, \( N \), is greater than 3.

Theorem 3. When the mean of the DKE exists,

\[
E[Z] = (a, b)^T + O(\sigma).
\]

Theorem 4. When the variance of the DKE exists,

\[
\text{var}(Z, Z) = r^2 \sigma^2 (S^T S)^{-1} + O(\sigma^3).
\]
The authors would like to acknowledge the contribution of Professor Barry Quinn whose many helpful comments have led directly to the improvement of this paper. We would also like to thank Dr. Peter Kootsookos for providing feedback on drafts of this paper.

5. Acknowledgement

The authors would like to acknowledge the contribution of Professor Barry Quinn whose many helpful comments have led directly to the improvement of this paper. We would also like to thank Dr. Peter Kootsookos for providing feedback on drafts of this paper.

6. References


