

Classification
Physics Abstracts
02.50

Recognition of the Form of Planar Objects by Means of Statistical Tests

Franz Streit

Section de Mathématiques de l'Université de Genève, Case postale 240, 1211 Genève 24, Switzerland

Résumé. — Dans cette étude on montre l'utilité des techniques statistiques pour reconnaître la forme commune d'objets plans à partir d'un système de ces objets soumis à des déformations aléatoires de leurs contours.

Abstract. — In this article the use of statistical technics to discover the underlying form of objects subject to random deformations is explained.

1. The Problem

In experimental sciences we frequently observe systems of objects of similar but not identical form which are dispersed in an Euclidean space \mathbb{R}^k , for instance in a plane \mathbb{R}^2 . Often it is reasonable to suppose that the individual objects of the system are obtained from an object of ideal form by deforming the contour of this object in a random way. It is then of interest to try to reconstruct the original ideal form based on a study of the individual deformed objects of the system. It is the goal of this paper to demonstrate the utility of simple statistical techniques to solve problems of this kind.

2. On a Model Proposed in a Book of D. and H. Stoyan

Of course there exist different ways to formalize the influence of random fluctuations in this context. In order to explain the main ideas I shall consider systems of planar objects and shall use a stochastic model proposed in the book of D. and H. Stoyan [6, p.110, f.(2.72)]. We suppose that the actually observed objects $O_j [j = 1, \dots, n]$ may be described by indication of a reference point P_j , which is an interior point of O_j (for instance the center of gravity of O_j), and the radius vector function R_j of its contour measured from the reference point and from a reference direction, which may be specified either by a fixed direction in \mathbb{R}^2 or by a reference point on the contour of O_j . $R_j(\varphi)$ is thus the length of the segment that emanates from P_j in φ -direction and ends on the contour of O_j . It is supposed that O_j is a non-empty compact set which is starlike, *i.e.* the above

mentioned segments are all subsets of O_j . We suppose furthermore that the forms K_1, \dots, K_n of these objects constitute a random sample, *i.e.* that K_1, \dots, K_n are independent and identically distributed random sets obtained by relocation of O_j to the origin of \mathbb{R}^2 , thus $K_j = (O_j)_{-P}$, where A_x designates the translate of A by x . Note that our interpretation of form includes an indication of the size of the object and of its orientation as in [5, pp. 35,36] and opposed to the use of this notion in [6, p.72]. Designating by r the radius vector function of the ideal form, the proposed model for the random deformation of the contour relates the quantities as follows:

$$R_i(\varphi) = r(\varphi) + Z_i(\varphi) \quad [0 \leq \varphi < 2\pi; i = 1, \dots, n].$$

We shall first discuss the situation for the case where we admit ellipses as forms. In standard orientation $r(\varphi)$ satisfies the main formula

$$r(\varphi) = (\cos^2(\varphi)/a^2 + \sin^2(\varphi)/b^2)^{-1/2} \quad [0 \leq \varphi \leq \pi/2]$$

together with the complementary symmetry formulae

$$\begin{aligned} r(\varphi) &= r(\pi - \varphi) & [\pi/2 < \varphi \leq \pi] \\ r(\varphi) &= r(\varphi - \pi) & [\pi < \varphi < 2\pi] \end{aligned}$$

with half-axes $a, b \geq 1$; a and b are unknown. The study of K_1, \dots, K_n is concretely carried out by measuring the radius vector on prescribed angles, for instance for $\varphi = 0^\circ, 90^\circ, 180^\circ$ and 270° . We suppose that the individual deformations $Z_1(l\pi/2), \dots, Z_n(l\pi/2)$, [$l = 0, 1, 2, 3$] at these angles form independent random samples from a population with density function

$$\begin{aligned} f_Z(z) &= 3[(2\pi)^{1/2}0.99865]^{-1} \exp[-9z^2/2] & [z \geq -1] \\ f_Z(z) &= 0 & [z < -1]. \end{aligned}$$

Thus the additive deformations follow a truncated normal distribution. The truncation is chosen in such a way that the radius vector of the resulting form does not take negative values.

3. Are the Object of Circular or of Non-Degenerated Elliptical Form?

Consider the following testing problem, which is of practical relevance in the context of the stochastic model described in the previous paragraph. We want to know, whether we need effectively to admit the class of ellipses or whether the more restricted class of circles would be sufficient for our purpose, *i.e.* we want to test the hypothesis

$$H_0 : a = b$$

versus the hypothesis

$$H_1 : a \neq b$$

where a and b are unknown, but $a, b \geq 1$. In such a case statistical methodology [3, 7, 8] recommends to use the generalized likelihood ratio test for the decision finding. The test-statistic is then calculated according to the formula

$$\Lambda_{\text{gen}} = [\sup_{a \geq 1, b \geq 1} (L(a, b; K_1, \dots, K_n))]^{-1} [\sup_{a=b=r \geq 1} (L(a, b; K_1, \dots, K_n))],$$

where L designates the likelihood-function, specified by

$$L(a, b; K_1, \dots, K_n) = \prod_{i=1}^n \prod_{l=0}^3 f_Z(Z_i(l\pi/2)) = \prod_{i=1}^n \prod_{l=0}^3 f_Z(R_i(l\pi/2) - r(l\pi/2; a, b)),$$

where the principal arguments of L are the unknown parameters of the model a and b and L is conditioned by the observed forms K_1, \dots, K_n insofar as known from the measurements $R_i(l\pi/2)$, $[i = 1, \dots, n; l = 0, 1, 2, 3]$.

For the values of the random variables, which are used in the evaluation of Λ_{gen} , we find

$$\begin{aligned} Z_i(l\pi/2) = R_i(l\pi/2) - r(l\pi/2 : a, b) &= R_i(l\pi/2) - a && [l = 0, 2] \\ &= R_i(l\pi/2) - b && [l = 1, 3] \quad [i = 1, \dots, n]. \end{aligned}$$

The numerator (resp. the denominator) of Λ_{gen} is obtained by replacing a and b by their common (resp. individual) maximum-likelihood estimators \hat{R} (resp. \hat{A} and \hat{B}). These estimators are obtained in the same way as in standard theory for normally distributed samples, but taking into account the size restrictions on a and b . Thus we find

$$\hat{R} = \max\{1, (4n)^{-1} \sum_{i=1}^n \sum_{l=0,1,2,3} R_i(l\pi/2)\}$$

and

$$\hat{A} = \max\{1, (2n)^{-1} \sum_{i=1}^n \sum_{l=0,2} R_i(l\pi/2)\}$$

$$\hat{B} = \max\{1, (2n)^{-1} \sum_{i=1}^n \sum_{l=1,3} R_i(l\pi/2)\}.$$

This leads to the explicit analytic expression of the test statistic given by

$$\begin{aligned} \Lambda_{\text{gen}} = \exp[9 \sum_{i=1}^n \sum_{l=0,2} (R_i(l\pi/2) - \hat{A})^2/2 + 9 \sum_{i=1}^n \sum_{l=1,3} (R_i(l\pi/2) - \hat{B})^2/2 \\ - 9 \sum_{i=1}^n \sum_{l=0,1,2,3} (R_i(l\pi/2) - \hat{R})^2/2] \end{aligned}$$

and for large values of n to the asymptotical critical region of size α

$$-2 \ln(\lambda_{\text{gen}}) > \chi_1^2(1 - \alpha). \tag{1}$$

In this context λ_{gen} designates the realized value of the statistic Λ_{gen} and $\chi_1^2(1 - \alpha)$ is the critical value of the χ^2 distribution of Pearson with one degree of freedom, which satisfies the relation

$$P(\chi_1^2 > \chi_1^2(1 - \alpha)) = \alpha.$$

For sufficiently large n the test is thus easy to carry out and leads to the rejection of H_0 if λ_{gen} satisfies the inequality (1); only a table of the critical values of the χ_1^2 -distribution is needed for the practical implementation of the procedure.

4. Generalizations and Extensions

Of course only a very special version of the problem of recognition of the underlying form has been treated so far and it is obvious that the same techniques may be used to resolve many similar problems.

4.1 OTHER CLASSES OF FORMS . — For instance we may replace the class of ellipses by a richer class of forms for which $r(0) = r(\pi) = a$ and $r(\pi/2) = r(3\pi/2) = b$. Such a class is for instance the class of superellipses [2], defined by the main formula

$$r(\varphi) = (\cos^c(\varphi)/a^c + \sin^c(\varphi)/b^c)^{-1/c} \quad [0 \leq \varphi \leq \pi/2] \quad [c > 0]$$

and the same complementary symmetry formulae as for the ellipses. The class admits convex forms (for $c \geq 1$) and non-convex forms (for $c < 1$).

For this class of forms the test of the hypothesis of equality of the half-axes a and b may be carried out in exactly the same way as previously described.

If one wants also to make inference on c , for instance if one wants to estimate c , it will be necessary to measure the radius vector function of the K_i 's at other angles than $0^\circ, 90^\circ, 180^\circ, 270^\circ$. For instance it may be advantageous to take additional measurements at the angles $45^\circ, 135^\circ, 225^\circ$ and 315° . We shall comment on this point later in Section 4.4.

4.2 RANDOM DEFORMATIONS GENERATED BY OTHER DISTRIBUTIONS . — Of course one can adopt probability laws for the random deformations $Z_i(l\pi/2)$ other than the truncated normal distribution. The procedure of generalized likelihood ratio test may still be useful, but the formulae for the maximum likelihood estimators and also the test statistic may change. One may also extend the techniques discussed to situations where the deformations at different angles are not independent. This could improve the applicability of the proposed methods, but leads to the problem of finding a sufficiently simple and generally meaningful incorporation of dependence relations into the model. Work on this is currently in progress.

4.3 OTHER DIMENSIONS OF THE SPACE . — Instead of considering systems of planar objects one may study systems of objects in other Euclidean spaces, for instance in three-dimensional spaces. One adapts then the formula for the ellipses (superellipses) by including a third coordinate. A practical context where such a set-up could be of interest is the situation, where one has to decide whether the sand grains found on a particular site have spherical or non-degenerate ellipsoidal form.

4.4 OTHER STOCHASTIC MODELS . — The additive stochastic model used so far may not be the only one to take into account. It may for instance be that the observed radius vectors are rather multiplicatively than additively related to the radius vectors of the common form as in the formula

$$R_i(\varphi) = Z_i(\varphi)r(\varphi) \quad [i = 1, \dots, n; 0 \leq \varphi < 2\pi]. \quad (2)$$

This type of model is simpler to handle mathematically, since in principle all the distributions concentrated on non-negative real values are now admitted for the specification of the probability law of the Z_i 's; we do not have to apply unusual distributions obtained by truncation in order to keep the R_i 's non-negative.

Suppose that we have already established that in the stochastic model with the relation (2) the class of superellipses satisfying $a = b$ is appropriate, but that c is still unknown and should be estimated. Suppose that the multiplicative deformation factors $Z_i(l\pi/4)$, [$i = 1, \dots, n$; $l = 1, 3, 5, 7$] form a random sample of size $4n$ from the exponential distribution with density function

$$f_Z(z) = \exp[-z] \quad [z > 0]$$

and

$$f_Z(z) = 0 \quad [z \leq 0].$$

We find

$$r(l\pi/4) = 2^{1/2-1/c} r \quad [l = 1, 3, 5, 7],$$

where $a = b = r$. Since $R_i(l\pi/4) = Z_i(l\pi/4)/r(l\pi/4)$, we find for the loglikelihood of the measurements which depend on c

$$L^* = -4n \ln(r2^{1/2-1/c}) - (r2^{1/2-1/c})^{-1} \sum_{l=1,3,5,7} \sum_{i=1}^n R_i(l\pi/4).$$

Thus

$$dL^*/dc = ([r2^{1/2-1/c}]^{-1}(-4n) + (r2^{1/2-1/c})^{-2} \sum_{l=1,3,5,7} \sum_{i=1}^n R_i(l\pi/4))r2^{1/2-1/c} \ln(2)c^{-2}.$$

Putting this expression equal to 0 yields the maximum likelihood equation

$$r2^{1/2-1/\hat{c}} = \bar{R}^* = (4n)^{-1} \sum_{l=1,3,5,7} \sum_{i=1}^n R_i(l\pi/4).$$

Therefore we obtain as estimator

$$\hat{c} = \max\{[\ln(2^{1/2}r/\bar{R}^*)]^{-1} \ln(2), 0\}.$$

4.5 OTHER APPROACHES . — Instead of treating the problems by statistical methodology, one can also use methods of morphology, like constructing and studying the skeleton of the individual objects [1, pp. 308–311] or treat the problems in a descriptive way using shape parameters (see e.g. [4] for the definition of such quantities).

References

[1] Coster M. and Chermant J.-L. Précis d'analyse d'images (les Éditions du CNRS, Paris, 1985).
 [2] Gardner M., Mathematical games: the "superellipse", a curve that lies between the ellipse and the rectangle, *Scient. Amer.* **213** (1965) 222-232.
 [3] Mood A.M. and Graybill F.A., Introduction to the Theory of Statistics (McGraw Hill, New York, 1963, 12.6).
 [4] Labouré M.-J., Jourlin M., Fillière I., Becker J.-M. and Frénéa F., Isoperimetric inequalities and shape parameters, *Acta Stereol.* **15** (1996) 65-70.
 [5] Sprent P., The mathematics of size and shape, *Biometrics* **28** (1972) 23-37.
 [6] Stoyan D. and Stoyan H., *Fraktale, Formen, Punktfelder. Methoden der Geometriestatistik* (Akademie-Verlag, Berlin, 1992) [English version (J. Wiley, New York, 1994)].
 [7] Wilks S., *Mathematical Statistics* (J. Wiley, New York, 1963, 13.3-13.8)
 [8] Witting H. and Noelle G., *Angewandte Mathematische Statistik* (Teubner, Stuttgart, 1970, 2.8).