STUDIA UNIV. BABEŞ–BOLYAI, INFORMATICA, Volume **XLVI**, Number 2, 2001

CIRCULAR FITTINGS FOR FINITE PLANAR POINT SETS

GHEORGHE CIMOCA

Abstract. Using the power of a point with respect to a circle, a (hopefully new) circular fitting for a finite point set in $\mathbb{R}^2$ is proposed. Finally, necessary and sufficient conditions for the collinearity and, respectively the concyclicity of n ($n \geq 2$, respectively $n \geq 3$) distinct points in $\mathbb{R}^2$ are given.

1. Introduction

Twenty five years ago, while the writing of [6] was in progress, we needed an algorithm to approximate a finite planar point set by an arc of a circle. The generated points were located on a certain area (specifically the foot) of a gear-tooth profile. At that moment we were not able to find any specific reference, but using the least squares method we quickly fitted a first model, and we later derived a second model as a minimax approximation problem [2]. The resulting algorithms were good enough for the specific purpose. However, the uniqueness of the solution was not assured because of the non-linearity of the normal equations of the first model, respectively the non-linear non-convex programming problem arisen in the second approach.

More recently, while designing a special disc milling cutter [1], we encountered the same approximation problem. This time, the uniqueness of the approximation circle was essential. Consequently, we returned to the old attempts and finally, by figuring out a new approach, we arrived at a simple and, simultaneously, very suitable solution. The result might be known, although we haven't found it mentioned anywhere, not even in the regression "Bible" [4].

In the subsequent paragraph we shall present both the non-linear fit models [2] and the new circular approximation technique for finite planar point sets.

2000 *Mathematics Subject Classification.* 65D10, 65J02.

1998 *CR Categories and Descriptors.* I.3.5[**Computational Geometry** and **Object Modelling**].

2. Circular fittings of finite point sets in $\mathbb{R}^2$

Let S be a finite set of points in the Euclidean plane $\mathbb{R}^2$. Let us assume that S contains at least 3 points, not all of them being collinear.

Our problem is to select a point $C(x, y) \in \mathbb{R}^2$ as the center of a circle, and a positive number z as its radius, so that this circle best approximates the considered point set S.

Related to our subject is the "minimum spanning circle" problem, in which the smallest circle that encloses a given set of points is sought [5]. The smallest enclosing circle is unique and is either the circumcircle of three extremal points of the convex hull of the given point set, $CH(S)$, or is defined by two of them as a diameter. Hence, the algorithm based on the previous remark examines only the extremal points of $CH(S)$.

On the other hand, the "largest empty circle" problem (i.e., finding the largest circle which contains no points of the set S) has neither a bounded nor a unique solution. In order to avoid unboundedness, a certain restriction regarding the center should be imposed. For instance, $C \in CH(S)$ is presented in [5] as a "necessary" condition, which is excessive and inaccurate. Restricting C so that it belongs to any arbitrarily given compact subset of $\mathbb{R}^2$ would suffice. Anyway, the uniqueness can't be guaranteed.

Such problems frequently arise in industrial engineering — e.g., the siting of emergency facilities (police stations, hospitals, &c.), the location of radio transmitters, the positioning of a source of pollution or of a new business that is not going to compete for territory with established rival outlets.

In contrast to the above-mentioned proximity (covers or gaps) problems, we shall consider all the points of S by means of their "distances" to the circle in question. The first two approximation models use the Euclidean distance from a point to a circle, measured on the normal direction, whereas for the third model the distance is replaced by the power of the point with respect to the circle. In fact, these quantities are estimates of the errors.

Let $S = \{P_i(a_i, b_i) \in \mathbb{R}^2 \ ; \ i = 1, \ \ldots, n\}$, where $3 \leq n < \infty$, $P_i \neq P_j$, for $i \neq j$, $i, j = 1, \ \ldots, n$, not all P_i being collinear.

Circular fit model I, using the least squares method. Minimize:

$$\sum_{i=1}^{n} [\sqrt{(x - a_i)^2 + (y - b_i)^2} - z]^2$$

in respect with the real variables x, y, z, where $0 \leq z \leq M$, M being a technological restriction.

The normal equations are:

$$\sum_{i=1}^{n}(x-a_i) - z\cdot\sum_{i=1}^{n}\frac{x-a_i}{\sqrt{(x-a_i)^2+(y-b_i)^2}} = 0,$$

$$\sum_{i=1}^{n}(y-b_i) - z\cdot\sum_{i=1}^{n}\frac{y-b_i}{\sqrt{(x-a_i)^2+(y-b_i)^2}} = 0,$$

$$n\cdot z - \sum_{i=1}^{n}\sqrt{(x-a_i)^2+(y-b_i)^2} = 0.$$

Substituting

$$z = \frac{1}{n}\cdot\sum_{j=1}^{n}\sqrt{(x-a_j)^2+(y-b_j)^2}$$

we obtain the following nonlinear system:

$$\sum_{i=1}^{n}\{(x-a_i)[n\cdot D_i(x,y) - \sum_{j=1}^{n}D_j(x,y)]\} = 0,$$

$$\sum_{i=1}^{n}\{(y-b_i)[n\cdot D_i(x,y) - \sum_{j=1}^{n}D_j(x,y)]\} = 0,$$

where $D_i^2(x,y) = (x-a_i)^2+(y-b_i)^2$.

Taking the barycenter of S as the initial approximation, the well-known Newton-Raphson algorithm [3] works quite well. Some remarks concerning the circle uniqueness appear in [2].

Circular fit model II, using the Chebyshev norm. Minimize:

$$\max\left\{\left|\sqrt{(x-a_i)^2+(y-b_i)^2} - z\right|, i=1,\dots,n\right\}$$

in respect with the real variables x, y, z, where $0 \le z \le M$.

Introducing a new variable u, the above problem can be reformulated as the following mathematical programming problem:

Minimize u
subject to:

$$\begin{aligned} x^2+y^2-z^2-u^2-2uz-2xa_i-2yb_i+a_i^2+b_i^2 &\le 0, \quad i=1,\dots,n,\\ x^2+y^2-z^2-u^2+2uz-2xa_i-2yb_i+a_i^2+b_i^2 &\ge 0, \quad i=1,\dots,n,\\ x\in\mathbb{R},\ y\in\mathbb{R},\ 0\le z\le M,\ 0\le u\le\varepsilon \end{aligned}$$

where ε is a maximum admissible error.

More details on this method, called the Megiddo method, can be found in [5].

Now let's remember that the number $\rho = d^2 - r^2$, where d is the distance from a point $P \in \mathbb{R}^2$ to the center C of a circle of radius r, denoted by (C, r), is called the power of the point P with respect to the circle. For any point lying outside the circle (C, r) we have the inequality $\rho > 0$, and for all points lying inside that circle we have $\rho < 0$. If $P \in (C, r)$ we obviously have $\rho = 0$.

Using the power of a point with respect to a circle we can derive:

Circular fit model III. Minimize:

$$\sum_{i=1}^{n} [(x - a_i)^2 + (y - b_i)^2 - z^2]^2,$$

in respect with the real variables x, y, z, where $0 < z \leq M$.

For this problem, the normal equations are:

$$\sum_{i=1}^{n} (x - a_i)[(x - a_i)^2 + (y - b_i)^2 - z^2] = 0,$$
$$\sum_{i=1}^{n} (y - b_i)[(x - a_i)^2 + (y - b_i)^2 - z^2] = 0,$$
$$\sum_{i=1}^{n} [(x - a_i)^2 + (y - b_i)^2] - nz^2 = 0.$$

Substituting

$$z^2 = \frac{1}{n} \cdot \sum_{j=1}^{n} [(x - a_j)^2 + (y - b_j)^2]$$

into the first two equations, we get the following system:

$$\sum_{i=1}^{n} (x - a_i)\{n[(x - a_i)^2 + (y - b_i)^2] - \sum_{j=1}^{n} [(x - a_j)^2 + (y - b_j)^2]\} = 0,$$
$$\sum_{i=1}^{n} (y - b_i)\{n[(x - a_i)^2 + (y - b_i)^2] - \sum_{j=1}^{n} [(x - a_j)^2 + (y - b_j)^2]\} = 0.$$

Apparently, this is a nonlinear system; yet, having performed all calculations, we finally obtain a linear system with the following unique solution:

$$x_0 = \frac{1}{2} \cdot \frac{B \cdot E - C \cdot D}{B^2 - A \cdot C} \tag{1}$$

$$y_0 = \frac{1}{2} \cdot \frac{B \cdot D - A \cdot E}{B^2 - A \cdot C} \tag{2}$$

where

$$A = \sum_{i=1}^{n} \sum_{j=1}^{n} (a_i - a_j)^2 = 2 \cdot \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (a_i - a_j)^2,$$

$$B = \sum_{i=1}^{n} \sum_{j=1}^{n} (a_i - a_j)(b_i - b_j) = 2 \cdot \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (a_i - a_j)(b_i - b_j),$$

$$C = \sum_{i=1}^{n} \sum_{j=1}^{n} (b_i - b_j)^2 = 2 \cdot \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (b_i - b_j)^2,$$

$$D = \sum_{i=1}^{n} \sum_{j=1}^{n} (a_i - a_j)(a_i^2 - a_j^2 + b_i^2 - b_j^2) = 2 \cdot \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (a_i - a_j)(a_i^2 - a_j^2 + b_i^2 - b_j^2),$$

$$E = \sum_{i=1}^{n} \sum_{j=1}^{n} (b_i - b_j)(a_i^2 - a_j^2 + b_i^2 - b_j^2) = 2 \cdot \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (b_i - b_j)(a_i^2 - a_j^2 + b_i^2 - b_j^2),$$

and the radius of the circle is:

$$z_0 = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} [(x_0 - a_i)^2 + (y_0 - b_i)^2]}.$$

3. Remarks and consequences

The fit model using the power of a point can be successfully applied also to conical (elliptic, hyperbolic or parabolic) fittings for finite planar point sets. Moreover, this model works very well in $\mathbb{R}^3$ to approximate a finite non coplanar point set by a spherical surface.

We'll end by mentioning the following obvious consequences:

Corollary 1. The points $P_i(a_i, b_i) \in \mathbb{R}^2, i = 1, \ldots, n$, where $2 \leq n < \infty$ are collinear if and only if:

$$\{\sum_{i=1}^{n} \sum_{j=1}^{n} (a_i - a_j)(b_i - b_j)\}^2 = \{\sum_{i=1}^{n} \sum_{j=1}^{n} (a_i - a_j)^2\} \cdot \{\sum_{i=1}^{n} \sum_{j=1}^{n} (b_i - b_j)^2\}.$$

The condition can be used for checking the (non)collinearity in problems related to certain geometrical data structures, such as the triangulations or the Voronoi diagrams [5].

Corollary 2. The points $P_i(a_i, b_i) \in \mathbb{R}^2, i = 1, \cdots, n$, where $3 \leq n < \infty$ are concyclic if and only if:

$$n \cdot \sum_{i=1}^{n} [(x_0 - a_i)^2 + (y_0 - b_i)^2]^2 = \{\sum_{i=1}^{n} [(x_0 - a_i)^2 + (y_0 - b_i)^2]\}^2$$

where $x_0, y_0 \in \mathbb{R}$ are given by (1) and (2), respectively.

We believe the above condition could be exploited in computational geometry as well.

References

[1] I. Bass, *General equations of gear cutting tool calculations; Gear Technology*, The Journal of Gear Manufacturing, 2 (1985), pp. 20–23.

[2] G. Cimoca, *Asupra aproximării prin arce de cerc a mulţimilor finite de puncte din plan*, Al II-lea Simpozion de informatică şi conducere, vol. II, C.T.C.E. Cluj-Napoca, 1977, pp. 104–106.

[3] W.S. Dorn, D.D. McCracken, *Numerical Methods with FORTRAN IV Case Studies*, Wiley, New York, 1972.

[4] W. Mendenhall, T. Sincich, *A Second Course in Statistics: Regression Analysis*, Prentice Hall, Upper Saddle River, New York, 1996.

[5] F.P. Preparata, M.I. Shamos, *Computational Geometry: An Introduction*, Springer Verlag, New York, Heidelberg, 1985.

[6] I.A. Stoica, G. Cimoca, *Interferenţa de sarcină a angrenajelor (Load Interference of Gears)*, Editura Dacia, Cluj-Napoca, 1978.

S.C. SIMBOLIC, str. Bacău nr. 3, 3400 Cluj-Napoca, ROMÂNIA, tel: +40-64-431.333, fax: +40-64-199.895, e-mail: ghcimoca@symbolic.com