IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 14, Issue 2 Ver. II (Mar. - Apr. 2018), PP 74-83
www.iosrjournals.org

The Algebraic Least Squares Fitting Of Ellipses

Abdellatif Bettayeb
Department of General Studies, Jubail Industrial College, Jubail Industrial City, Kingdom of Saudi Arabia
Corresponding Author: Abdellatif Bettayeb

Abstract: *Fitting ellipses to a set of given points in the plane is a problem that arises in many application areas, e.g. computer graphics [9], [3], coordinate metrology [2], petroleum engineering [8]. In this paper, we present several algorithms which the ellipse for which the sum of the squares of the distances to the given points is minimal. These algorithms are compared with classical and iterative methods. Ellipses is represented algebraically i.e. by an equation of the form F(x) = 0. The algorithm computes a continuous function closely approximating the ellipses, for which the sum of the squares to the given set of points is minimized. We will look particularly at one method, by giving examples and using Matlab to solve these problems and then compares the efficiency of them.*

Date of Submission: 11-04-2018 Date of acceptance: 26-04-2018

I. Introduction

Let a relationship between variables x and y be given by $f(x,y;p)=0$, where $p \in R^n$ is a vector of parameters. For example, this could be an ellipse or any conic in the x,y plane.

Let data points $(x_i, y_i),\ i=1,\ldots,m$ be given. Then ideally we wish to choose p so that

$$f(x_i, y_i; p) = 0, \quad i = 1,..., m.$$

However, this is unlikely to be possible, so we need some other ways of choosing p.

Different methods are available, and are considered in the next sections. One way of choosing p is called "Algebraic Fitting ", uses the implicit form of the ellipse, and takes the form of a constrained least square problem. This algorithm was extensively discussed by Varah in [9] and Golub, Gander and Strebel in [3].

In the other sections we introduces different numerical different numerical examples, with relevant figures and results. Then we compare all the methods studied in this paper using different examples.

II. Algebraic Fitting

Given the implicit function

$$f(x, y; p) = 0,$$

and the data points (x_i; y_i); i = 1; : . : ; m in the plane, we want to minimise the standard least square functional

$$\phi(p) = \sum_i f^2(x_i, y_i; p) \tag{2.1}$$

Generally, this results in a nonlinear minimisation problem, but in the important case that f is linear in p, we obtain a linear least squares problem of the form

$$\min_p \phi(p) = \min_p \|Sp\|_2^2, \tag{2.2}$$

such that the i-th component of Sp is

$$f(x_i, y_i; p), \qquad i = 1,2,..., m,$$

where S is an m x n matrix. We assume m > n and p is an n x 1 vector (parameter).Clearly p =0 is a trivial solution. Therefore to complete the problem, we must add some normalization condition on p, and here we have four different possibilities.

These give rise to constrained optimization problems and therefore we need to introduce some analysis of problems of this kind. This is considered next.

III. The constrained optimization problem

The structure of most constrained optimization problems is essentially contained in the following:

$$Minimize \quad f(x) \qquad ,x \in R^n \qquad (3.1)$$
$$Subject \quad to \quad c_i(x) = 0, \quad i \in E$$
$$c_i(x) \geq 0 \ \ ,i \in I,$$

where f(x) is objective function, but there are additional constraint functions $c_i(x),\ i = 1,2,...,\ k,$, where k is the number of indices in EUI, and E is the index set of equations or equality constraints in the problem, I is the set of inequality constraints, and both of these sets are finite. More general constraints can be put into this form: for example $c_i(x) \leq b$ becomes $b - c_i(x) \geq 0.$. If any point x satisfies all constraints in (3.1), it is said to be a feasible point and the set of all such points is referred to as the feasible region R (see [2] page 139-140).

3.1. SVD Constraint

$$\min_p \phi \quad subject\ to \quad \|p\|^2 = 1, \qquad (3.2)$$

where $\|.\|$ is the l_2 norm. We have the main problem as

$$\min_p \|Sp\|^2 \quad subject\ to \|p\|^2 = 1. \qquad (3.3)$$

The quantity $\|Sp\|^2$ isminimised at the same p that minimises $\|Sp\|$, and $\|Sp\|_2^2$ is easier to study. Observe that:

$$\|Sp\|_2^2 = (Sp)^T(Sp) = p^T S^T Sp = p^T(S^T S)p.$$

Also, S^TS is a symmetric matrix, since $(S^TS)^T = S^TS^{T\,T} = S^TS$.
So the problem now is to minimize the quadratic form $p^T(S^TS)p$ subject to the constraint $\|p\|^2 = 1$.
Let A be a symmetric matrix, and define m and M as:

$$m = \min x^T Ax : \|x\| = 1 \qquad M = \max x^T Ax : \|x\| = 1$$

Then by Theorem 6 page 421 from [5], M is the largest eigenvalue λ_1 of A and m is the smallest eigenvalue of A. The value of $x^TAx$is M when x is a unit eigenvector v_1 corresponding to M. The value of $x^T Ax$ is m when x is a unit eigenvector corresponding to m. Thus the minimum value is the eigenvalue corresponding to the smallest eigenvector v_n.
If S is an m n matrix, then $S^T S$ is symmetric as we saw before, and can be orthogonally diagonalized. Let $v_{1,...,}v_n$ be an orthogonal basis for R^n consisting of eigenvectors of $S^T S$, and let $\lambda_1,..., \lambda_n$ be the associated eigenvalues of S^TS.
Then, for $1 \leq i \leq n,$,

$$\|Sp_i\|^2 = (Sp_i)^T(Sp_i) = p_i^T S^T p_i = p_i^T(S^T S)p_i = p_i^T(\lambda_i p_i),$$

Since p_i is an eigenvector of $S^T S$

$$\|Sp_i\|^2 = \lambda_i, \qquad (3.4)$$

since p_i is a unit vector. So the eigenvalues of S^TS are all nonnegative. By numbering, if necessary, we may assume that the eigenvalues are arranged so that

$$\lambda_1 \geq \lambda_2 \geq ....\ \lambda_n \geq 0.$$

The singular values of S are the square roots of the eigenvalues of S^TS, denotedby $\sigma_1, \sigma_2, ..., \sigma_n$, and they are arranged in decreasing order. That is $\sigma_i = \sqrt{\lambda_i}$ for $1 \le i \le n.$

By the equation (3.4), the singular values of S are the lengths of the vectors $Sp_1, ..., Sp_n$ (see [6] page 428-430).

For $\|p\| = 1$, the minimum value of $\|Sp\|$ is $\|Sp_n\| = \sigma_n = \sqrt{\lambda_n}$.

We can write S in the form of a singular value decomposition (SVD) of S, say

$$S = U \sum V^T,$$

Where

$$\sum = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix},$$

$$D = \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r \end{bmatrix},$$

here is $\sum$ an m x n matrix, where the diagonal entries in D are the first r singular values of S, $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r >> 0,$, and U is an m x m orthogonal matrix and an n x n orthogonal matrix V.

Then we can summarise this, by saying the problem is equivalent to finding the right singular vector $p = V(:, n)$ associated with the smallest singular value of S.

Remark: The SVD method gives a good fit when $\sigma_n << 1.$. The singular vector $p = V(:, n)$ corresponding to σ_n is assumed unique.

3.2Quadratic constraint

$$\min_p \|Sp\|_2^2 \quad subject\ to\ \|Bp\|_2^2 = 1, \text{(3.5)}$$

whereB is a general k xn matrix.

We can find the solution p using the generalised SVD of S and B, or by solving a generalised eigenvalue problem involving S^TS and B^TB. However, we take here the special case when B = [O I], and we show in the following discussion how to find the solution p. For instance we take the case of ellipse with parameter $p = (a_{11}, 2a_{12}, a_{22}, b_1, b_2, c)^T$ with $\|p\| = 1.$.

If we define vectors $v = (b_1, b_2, c)^T$, $w = (a_{11}, \sqrt{2}a_{12}, a_{22})^T$ then $p = (v; w);$ and the coefficient matrix

$$S_{m\times n} = \begin{bmatrix} x_1 & y_1 & 1 & x_1^2 & \sqrt{2}x_1y_1 & y_1^2 \\ \vdots & & & & & \vdots \\ x_m & y_m & 1 & x_m^2 & \sqrt{2}x_my_m & y_m^2 \end{bmatrix},$$

theBookstein constraint i.e. the quadratic constraint may be written as $\|w\| = 1,$ and we have the system:

$$s\begin{bmatrix} v \\ w \end{bmatrix} = 0.$$

The QR decomposition of S leads to the equivalent system

$$\begin{bmatrix} R_1 & R_2 \\ 0 & R_3 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = 0,$$

and it is equivalent to

$$R_1 v + R_2 w = 0,$$

$$R_3 w = 0.$$

Which may be solved in the following steps:

1. $R_3 w = 0, \quad \|w\| = 1.$

Using the singular value decomposition of $R_3 = U \sum V^T, w$ is the singular vector corresponding to the smallest singular value of R_3.

2. Finally find v from: $R_1 v + R_2 w = 0,$ and this is equivalent to

$R_1 v = -R_2 w,$ and then: $v = -R_1^{-1} R_2 w.$

Thus

$$p = \begin{bmatrix} v \\ w \end{bmatrix} = 0.$$

3.3. Ellipse constraint

$$\min_p \|Sp\|^2 \quad \textit{subject to} \quad p^T C p = 1, \qquad (3.8)$$

Where C is

$$C = \begin{bmatrix} 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

We can reduce thisproblem to a generalised eigenvalue problem:

$$(C - \mu\, S^T S)\, p = 0. \qquad (3.9)$$

For the ellipse, the solution p is the eigenvector corresponding to the unique positive eigenvalue in the ellipse case, or the biggest absolute value of the eigenvalue .

Assuming S is nonsingular, this generalised eigenvalue problem has the same number of positive eigenvalues as C does, namely one.

More generally, the p takes the value of the eigenvector corresponding to the largest nonnegative absolute value of the previous problem as follows: If we apply the first condition of Lagrange multiplier (see [2]), we get

$$L(x, \lambda) = p^T S^T S p - \lambda (p^T C p - 1).$$

Differentiate w.r.t. to p:

Differentiate w.r.t to λ :

$$\nabla_p = 2S^T Sp - 2\lambda Cp = 0.$$

$$\nabla_\lambda = p^T Cp - 1 = 0.$$

Then we get the problem as above

$$2S^T Sp - 2\lambda Cp = 0 \Rightarrow (S^T S - \lambda C)p = 0$$

$$\Rightarrow S^T Sp = \lambda Cp \Rightarrow \left(\frac{1}{\lambda}\right)p = \left(S^T S\right)^{-1} Cp$$

$$\Rightarrow \left(\frac{1}{\lambda}\right)\left(S^T S\right)p = Cp \Rightarrow$$

$$\left(C - \left(\frac{1}{\lambda}\right)S^T S\right)p = 0 \Rightarrow Cp = \left(\frac{1}{\lambda}\right)\left(S^T S\right)p.$$

Then

$$1 = p^T Cp = \left(\frac{1}{\lambda}\right)p^T S^T Sp > 0,$$

thus

$$\|Sp\|^2 = p^T S^T Sp = \lambda,$$

and λ must be positive because $\|Sp\|^2 \geq 0,$ and to get a nontrivial solution

λ must be strictly positive. Numerically we found always λ corresponds to the n-2 column of (in MATLAB notation we put D instead of $\sum$) and the

solution p corresponds to the n-2 vector v of the matrix V (by MATLAB) such that $[V, D] = eig(M),$ where $M = (S^T S)^{-1} C$ because all the previous eigenvalues in the case of this matrix C arezeros. Then we want the

smallest positive value λ , which corresponds to the largest absolute value of

$\left(\frac{1}{\lambda}\right)$ of $\left(S^T S\right)^{-1} C.$

Remark:

Usually, for all the constraints given, the solutions obtained are very different, and lead to very different figures in general. In the case when the data are very nearly fitted by the implicit functions (i.e. where the residuals are small and $\sigma_n << 1$), the solutions obtained are quite similar. With the SVD for an ellipse, this is not well fitted because the matrix A, although positive definite $A = [a_{11}\ a_{12}; a_{12}\ a_{22}]$, is very nearly singular. We mention here that depending on the data, the generated matrix A may be indefinite or singular [9].

And now we guarantee the generation of a positive definite matrix A, and hence of an ellipse (or circle) rather than another conic section by the fourth constraint (Ellipse constraint).

IV. Algebraic Fitting of ellipses

Consider the implicit function of a general ellipse in the plane, which can expressed as

$$f(x, y; p) = X^T AX + b^T X + C,$$

where

$$X = \begin{bmatrix} x \\ y \end{bmatrix}, y = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix},$$

and p is defined as

$$p = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{22} \\ b_1 \\ b_2 \\ c \end{bmatrix}.$$

Then

$$\phi(p) = \|Sp\|^2,$$

where S is m x 6 matrix

$$S = \begin{bmatrix} x_1^2 & 2x_1y_1 & y_1^2 & x_1 & y_1 & 1 \\ \vdots & & & & & \vdots \\ x_m^2 & 2x_my_m & y_m^2 & x_m & y_m & 1 \end{bmatrix},$$

we can put

$$\|Sp\|_2^2 = (Sp)^T (Sp) = p^T S^T Sp.$$

The algebraic residual can be defined as

$$res = \|Sp\|_2^2 \Big/ \|p\|_2^2.$$

If we add to the problem, the ellipse constraint, we should have this form:

$$\min_p p^T S^T Sp \quad subject \quad to \quad p^T Cp = 1,$$

where C was defined earlier. We can guarantee A's positive definiteness by this method i.e. ellipse constraint, because A is positive definite if

$p(1)p(3)-p(2)^2>0 \Leftrightarrow p^TCp>0.$

This problem can be reduced to the generalised eigenvalue problem:

$(C-\mu S^TSp)=0,$

and this is equivalent to
$Cp=\mu(S^TS)p,$

where $(C-\mu S^TS)$, C and S^TS have the same order. The eigenvalues of $C-\mu(S^TS)$ are the eigenvalues of $(S^TS)^{-1}C.$ And they are also the eigenvalues of $CB^{-1}=B(B^{-1}C)B^{-1},$ and $B^{-1}C=(S^TS)C,$ where (see [7] and [6]).

If B is nonsingular i.e. $(\det(A)\neq 0)\leftrightarrow B^{-1}Cp=\mu p.$

The solution p is the eigenvector corresponding to the unique positive eigenvalue in the case of ellipses.

After solving the algebraicfitting for p, we must construct the ellipses by converting its algebraic form to the parametric representation:

$X=Q\overline{X}+z,$

Where

$$z=\begin{bmatrix}x_c\\ y_c\end{bmatrix}, \overline{X}=\begin{bmatrix}\beta\cos\theta\\ q\sin\theta\end{bmatrix},$$

By substituting X in the equation

$X^TAX+b^TX+c=0.$

We get
$\overline{X}(Q^TAQ)\overline{X}+(2z^TA+b^T)Q\overline{X}+(z^TAz+b^Tz+c)=0,$

We put:
$\overline{A}=(Q^TAQ) \quad \overline{b}^T=(2z^TA+b^T)Q \quad \overline{c}=(z^TAz+b^Tz+c),$
And the equation will be

$\overline{X}^T\overline{A}\overline{X}+\overline{b}^T\overline{X}+\overline{c}=0.$

Choose Q so that: $\overline{A}=diag(\lambda_1,\lambda_2),$ where λ_1,λ_2 are the eigenvalues of A.
Choose z so that:

$\overline{b}=0\Rightarrow(2z^TA+b^T)Q=0\Rightarrow 2z^TA=-b^T\Rightarrow z^T=-b^T/2A.$

If $\lambda_1.\lambda_2>0\Rightarrow$ the conic is an ellipse with

$$z = \begin{bmatrix} x_c \\ y_c \end{bmatrix},$$

and $\beta = \sqrt{-\bar{c}/\lambda_1}$, $q = \sqrt{-\bar{c}/\lambda_2}$. A and $\bar{A}$ have the same (real) eigenvalues λ_1, λ_2.
Choose $Q = Vl$ where Vl is the matrix contains the eigenvectors of A corresponding to λ_1, λ_2 such that: $Vl^T Vl = I$.

Thus the parametric form of the ellipse is:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \end{bmatrix} + Q \begin{bmatrix} \beta\cos\theta \\ q\sin\theta \end{bmatrix}.$$

This is the case of Ellipse constraint, where θ varies from 0 to 2π.

4.1The General algorithm
Step 1: Calculate S.

Step 2:
Calculate the eigenvalues of $B^{-1}C$, where $B = S^T S$ and choose the smallest positive one i.e. (μ) or the biggest absolute value of λ.

Step 3:
The solution (p) would be the eigenvector corresponding to ( ) which is equivalent to the biggest positive value of such that $\mu = \frac{1}{\lambda}$.

V. Examples

With different set of points, we applied here the algebraic method for an ellipse. Matlab was used here, because it is easy to implement against another package or language like Fortran, and we save a lot of time too. Also we used the algebraic t with different constraints.

5.1Fitting Ellipses
Consider the $Sp\ddot{a}th$ data set given from [4] in Table (4.1) which is used for all examples of ellipses.

x y	8 1	3 6	2 3	7 7	6 1	6 10	4 0

5.2Example 1: Algebraic method with Ellipse Algorithm
Figure 1 shows the ellipse generated from the data using the ellipse constraint.
And the residual is $\text{res} = 2.727304\text{e}^{-01}$. The minimum of $\phi == 7.438188\text{e}^{-02}$.
The center is z = [5:2604 4:4887]^T .

The best fitobtained is the ellipse

$$x = 5.2604 + 2.9542\cos\theta$$

$$y = 4.4887 + 4.8224\sin\theta.$$

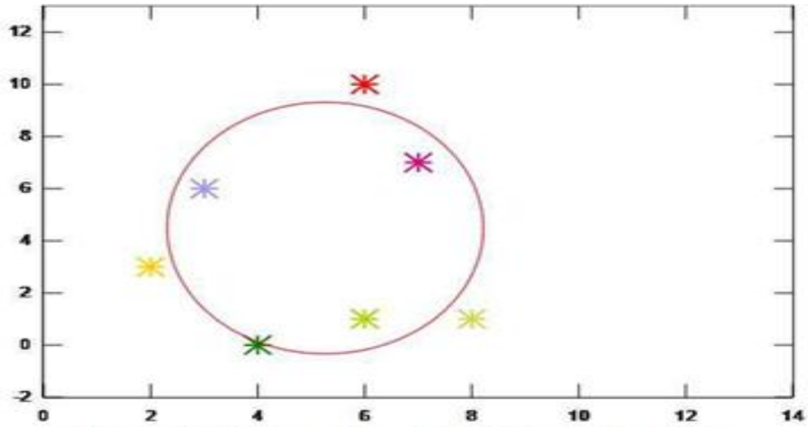

Figure 1: Ellipse fits with Ellipse Algorithm

5.3Example 2: Algebraic method with SVDAlgorithm

Figure 2 shows the ellipse generated from the data using SVD method.
The ellipse obtained is

$$x = 4.9293 + 2.5442 \cos\theta,$$

$$y = 5.1726 + 6.3990 \sin\theta.$$

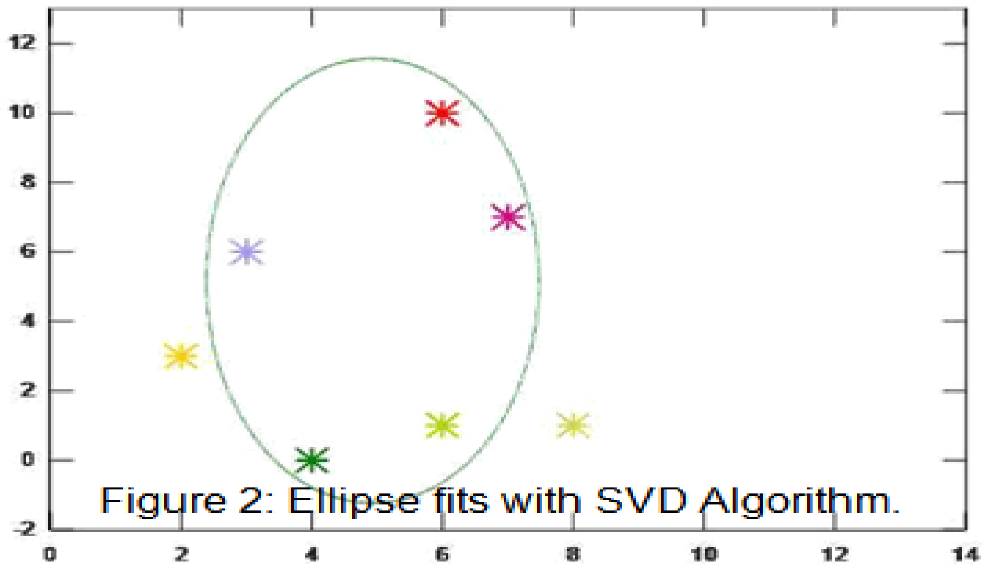

Figure 2: Ellipse fits with SVD Algorithm.

5.4Example 3: Algebraic method with Quadratic Algorithm

We have Figure 3 shows the ellipse generated from the data using the Quadratic Algorithm method.

$$\text{residual} = 2.382263\text{e}^{-01} \qquad \phi = 4.377026\text{e}^{+01}.$$

The ellipse generated is

$$x = 5.0270 + 2.5794 \cos\theta,$$

$$y = 4.5461 + 5.8333 \sin\theta.$$

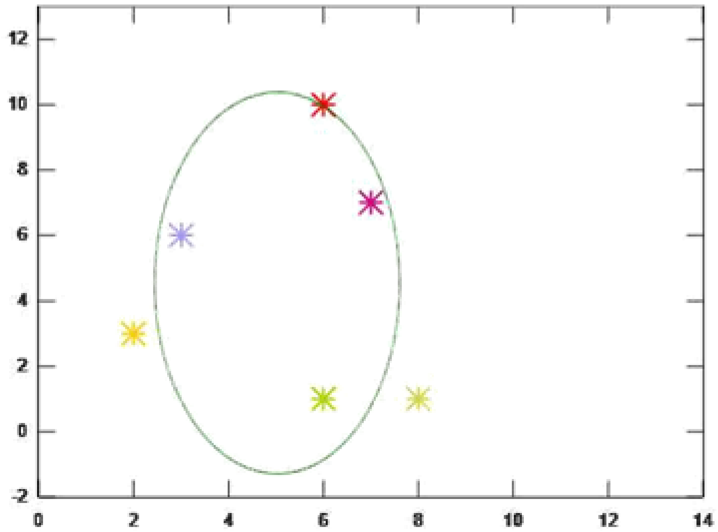

Figure 3: Ellipse fits with Quadratic Algorithm.

Bibliography

[1] Dixon, L. C. W.,Spedicato, E. and Szego, G. P., Nonlinear Optimization: theory and algorithms, Birkhauser Boston, 1980. ISBN 3-7643-3020-1. pp. 91-102.
[2] Fletcher, R. G., Practical Methods of Optimization, John Wiley and Sons, 1995. ISBN 0 471 91547 5 pp. 111-136.
[3] Gander, W., Golub, G. H. and Strebel, R., Fitting of circles and ellipses: least square solution,BIT, 34(1994), pp. 556-577.
[4] Huffel Sabine Van, Recent Advances in total least squares techniques and errors in variables modeling, " Orthogonal Least Squares Fitting by Conic Sections " by HelmuthSpath pp. 259-264, Library of Congress, USA, 1997. ISBN 0-89871-393-5.
[5] Lay, D. C., Linear Algebra and its applications, Addison-Wesley, 1994. ISBN 0-201-52031-1.
[6] Ortega, J. M., Matrix Theory, Plenum Press, 1987. ISBN 0 306 42433 9.
[7] Stewart, G. W., Introduction to Matrix Computations, Academic Press, INC, 1973.
[8] The MathWorks, Inc. MATLAB REFERENCE GUIDE, 1992. PP. 205.
[9] Varah, J. M., Least Squares Data Fitting with implicit functions, BIT, 36 (1996), pp. 842 - 854.

Abdellatif Bettayeb "The Algebraic Least Squares Fitting Of Ellipses. IOSR Journal of Mathematics (IOSR-JM) 14.2 (2018) PP: 74-83.