LEAST SQUARES FITTING OF ELLIPSOID USING ORTHOGONAL DISTANCES

Adequação do elipsóide usando distâncias ortogonais com mínimos quadrados

SEBAHATTIN BEKTAS
OndokuzMayis University, Faculty of Engineering, Geomatics Engineering, 55139 Samsun, sbektas@omu.edu.tr

Abstract:
In this paper, we present techniques for ellipsoid fitting which are based on minimizing the sum of the squares of the geometric distances between the data and the ellipsoid. The literature often uses “orthogonal fitting” in place of “geometric fitting” or “best-fit”. For many different purposes, the best-fit ellipsoid fitting to a set of points is required. The problem of fitting ellipsoid is encountered frequently in theimage processing, face recognition, computer games, geodesy etc. Today, increasing GPS and satellite measurements precision will allow ustodetermine amore realistic Earth ellipsoid. Several studies have shown that the Earth, other planets, natural satellites, asteroids and comets can be modeled as triaxial ellipsoids Burša and Šima (1980), Iz et al (2011). Determining the reference ellipsoid for the Earth is an important ellipsoid fitting application, because all geodetic calculations are performed on the reference ellipsoid. Algebraic fitting methods solve the linear least squares (LS) problem, and are relatively straightforward and fast. Fitting orthogonal ellipsoid is a difficult issue. Usually, it is impossible to reach a solution with classic LS algorithms. Because they are often faced with the problem of convergence. Therefore, it is necessary to use special algorithms e.g. nonlinear least square algorithms. We propose to use geometric fitting as opposed to algebraic fitting. This is computationally more intensive, but it provides scope for placing visually apparent constraints on ellipsoid parameter estimation and is free from curvature bias Ray and Srivastava (2008).

Keywords: Fitting Ellipsoid; Orthogonal Fitting; Algebraic Fitting; Nonlinear Least Square Problem.

Resumo: Neste trabalho, apresentamos a adequação do elipsóide com base na soma dos mínimos quadrados das distâncias geométricas entre os dados e o elipsóide. A literatura muitas vezes aborda a “adequação ortogonal” em lugar da “adequação geométrica” na “melhor adequação”. Para muitos propósitos diferentes, a melhor adequação do elipsóide para o conjunto de pontos é necessária. O problema da adequação do elipsóide é encontrado frequentemente no processamento de imagens, reconhecimento de superfícies, jogos de computador, geodésia, etc. Hoje, o aumento da precisão das medidas GPS e de satélites permite que se determine um elipsóide da Terra mais realista. Diversos estudos, tem demonstrado que a Terra, outros planetas, satélites naturais, asteróides e cometas podem ser modelados com o elipsóides triaxiais, Bursa e Sima (1980), Iz et al (2011). Determinar o
1. INTRODUCTION

Fitting an ellipsoid to an arbitrary set of points is a problem of fundamental importance in many wide fields of applied science ranging from astronomy, geodesy, digital image processing and robotics to metrology etc. Ellipsoids, though a bit simple in representing 3D shapes in general, are the only bounded and centric quadrics that can provide information of center and orientation of an object. Fitting ellipsoid has been discussed widely and some excellent work has been done in literature. However, most of these fitting techniques are algebraic fitting, but not orthogonal fitting. Various “least-squares” fitting approaches have been formulated over the years Zhang (1997), but they all fall into two categories; (1) algebraic methods, which are extensively used due to their linear nature, simplicity and computationally efficiency, and (2) geometric methods that solve a nonlinear problem Ray and Srivastava (2008).

We could not find enough studies with numerical examples in the literature. Turner et al (1999) gave a numerical application, but the application’s data are not given Turner et al (1999). No other comparable orthogonal fitting ellipsoid application could be found in literature. Against this background, the purpose of the study is to give an orthogonal fitting ellipsoid with numerical examples. In this article, we demonstrate that the geometric fitting approach, provides a more robust alternative than algebraic fitting approach - although it is computationally more intensive.

The paper has eight parts. First, the basic ellipsoid will introduce some mathematical equationsto explain the concepts. Then, it reviews the extended literature relevant to ellipsoid fitting. And we discussed in this research which estimators is used. Next, comes the part which deals with algebraic fitting, orthogonal fitting and numerical example. You will find ellipsoid fitting application based on both \( l_1 \)-norm and \( l_2 \)-norm methods. The paper concludes with a discussion of theoretical and managerial implications and directions for further research.

1.1 Ellipsoid

An ellipsoid is a closed quadric surface that is analogue of an ellipse. Ellipsoid has three different axes \((a_x>a_y>b)\) in Figure 1. Mathematical literature often uses “ellipsoid” in place of “Triaxial ellipsoid or General ellipsoid”. Scientific literature (particularly geodesy) often uses
“ellipsoid” in place of “biaxial ellipsoid, rotational ellipsoid or ellipsoid revolution”. Older literature uses ‘spheroid’ in place of rotational ellipsoid. The standard equation of an ellipsoid centered at the origin of a cartesian coordinate system and aligned with the axes is shown with this formula:

\[
\frac{x^2}{a_x^2} + \frac{y^2}{a_y^2} + \frac{z^2}{b^2} = 1
\]  

(1)

Although ellipsoid equation is quite simple and smooth, computations are quite difficult on the ellipsoid. The main reason for this difficulty is the lack of symmetry. Generally, an ellipsoid is defined with 9 parameters. These parameters are; 3 coordinates of center \((X_0, Y_0, Z_0)\), 3 semi-axes \((a_x, a_y, b)\) and 3 rotational angles \((\varepsilon, \psi, \omega)\) which represent rotations around \(x\)-, \(y\)- and \(z\)-axes respectively in Figure 2. These angles control the orientation of the ellipsoid.

\[R_1, R_2, R_3\] are plane rotation matrices

\[
R_1(\varepsilon) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varepsilon & \sin \varepsilon \\ 0 & -\sin \varepsilon & \cos \varepsilon \end{bmatrix}, \quad R_2(\psi) = \begin{bmatrix} \cos \psi & 0 & -\sin \psi \\ 0 & 1 & 0 \\ \sin \psi & 0 & \cos \psi \end{bmatrix}, \quad R_3(\omega) = \begin{bmatrix} \cos \omega & \sin \omega & 0 \\ -\sin \omega & \cos \omega & 0 \\ 0 & 0 & 1 \end{bmatrix}
\]

(2)

R-rotation matrix is obtained from \(R_1, R_2, R_3\) by multiplying the reverse order

\[
R = R_3(\omega)R_2(\psi)R_1(\varepsilon)
\]

\[
R = \begin{bmatrix} \cos \psi \cos \omega & \cos \varepsilon \sin \omega + \sin \varepsilon \sin \psi \cos \omega & \sin \varepsilon \sin \omega - \cos \varepsilon \sin \psi \cos \omega \\ -\sin \psi \sin \omega & \cos \varepsilon \cos \omega - \sin \varepsilon \sin \psi \sin \omega & \sin \varepsilon \cos \omega + \cos \varepsilon \sin \psi \sin \omega \\ \sin \varepsilon & -\sin \varepsilon \cos \psi & \cos \varepsilon \cos \psi \end{bmatrix}
\]

(3)
Figure 2: Shifted - oriented ellipsoid

2. FITTING ELLIPSOID

For the solution of the fitting problem, the linear or linearized relationship, written between the given data points and unknown parameters (one equation per data points), consists of equations, including unknown parameters.

\[ A_{n \times u} \cdot \delta x = l \]

Here, \( A \) is design matrix, \( \delta x \) is unknown parameters, \( l \) is measurements vector or data points, \( n \) is number of given data point (or measurements), \( u \) is number of unknown parameter, \( f = n - u \) : degree of freedom

-If \( f = 0 \) there is only one (exact) solution, algebraic solution
-If \( f < 0 \) there is no solution. The solution can be found with based on the extra constraint
-If \( f > 0 \) is most commonly encountered situation. The given data points (or measurements), which are much greater than the required number cause discrepancy, and in this case, the solution is not unique. There is an over determined system. Because \( n > u \), in other words the number of equations is greater than the number of unknowns.

The system of linear equations (3) must be solved. Therefore, this system must be consistent with the range of design matrix, and design matrix extended with constant terms, must be equal, so that \( \text{rang}(A) = \text{rang}(A:l) \); whereas, the system of (3) is inconsistent, because \( \delta x \) unknown parameters that provide (3), can not be calculated. In this case, \( \text{rang}(A) \leq u \). The extended matrix with \( l \) measurements \( \text{rang}(A:l) \) is generally more than \( \text{rang}(A) \). There is no solution of inconsistent equations, and only the approximate solution of the system can be derived. The equation system with approximate solution is calculated by adding \( \varepsilon \) residuals (or corrections) at the right side of (3).

Depending on the choice of $\varepsilon$ residuals vector, infinite solutions can be obtained. The unique solution can be derived only according to an estimator (objective function). For example, the LS always give an unique solution Bektas and Sisman (2010). Here, the question of which estimation method to use comes to mind?

3. WHICH ESTIMATOR SHOULD BE USED?

It is hoped that the residuals will be small. The more suitable estimation method is one that creates smaller residuals. It is seen that usually the objective functions are formed based on the minimization of corrections or a function of corrections. There are numerous estimators, some of these are $l_1$-norm, $l_2$-norm,$l_p$-norm, Fair, Huber, Cauchy, German-McClure, Welsch and Tukey. Two estimator methods come to the forefront. The most used estimators are shown below:

(i) $[\varepsilon] = \min. (l_2\text{-norm})$ Least Squares Method (LSM)
(ii) $[I\varepsilon] = \min. (l_1\text{-norm})$ Least Absolute Values Method (LAVM).

3.1. The Comparison of $l_1$ and $l_2$-Norm Methods

The solution of the $l_2$-norm method is always unique, and this solution is easily calculated. The $l_2$-norm method is widely used in parameters estimation. The $l_2$-norm method has indisputable superiority in parameter estimation.

The disadvantages of the $l_2$-norm method are that it is affected by outlying (gross errors) and it distributes to the sensitivity measurements. In this case, ellipsoid fitting is a very nice application.

With least-squares techniques, even one or two outliers in a large set can wreak havoc! Outlying data give an effect so strong in the minimization that the parameters thus estimated by those outlying data are distorted. Numerous studies have been conducted, which clearly show that least-squares estimators are vulnerable to the violation of these assumptions. Sometimes, even when the data contains only one bad measurement, $l_2$-norm method estimates may be completely perturbed Zhang (1997).

The solution of the $l_1$-norm method is not always unique, and there may be several solutions. Also, the solution of the $l_1$-norm method is not generally obtained directly, but iteratively calculations are made. Therefore, the solution is not easily calculated like in the $l_2$-norm method. Notwithstanding, when the computational tools, computer capacity and speed are considered, the difficulty of calculations are eliminated. The advantages of the $l_1$-norm method are non-sensitivity against measurements, including gross errors, and the solution is not or is little affected by these measurements.

The author of this study proposed and used the $l_2$-norm method in the solution of parameter estimation (optimization problems, adjustment calculus), after the measurement group cleaned up gross and systematic errors using the $l_1$-norm method. For further information see Bektas and Sisman (2010).

4. ALGEBRAIC ELLIPSOID FITTING METHODS
The general equation of an ellipsoid is given as

\[ A'x^2 + B'y^2 + C'z^2 + 2D'xy + 2E'xz + 2F'yz + 2G'x + 2H'y + 2I'z + K' = 0 \]  

(5) contains ten parameters. In fact, nine of those ten parameters are independent. For example, if all the coefficients in this equation multiply by \((-1/K')\), we get a new equation which contains nine unknown parameters, and its constant term will be equal to \("-1\".

\[ Ax^2 + By^2 + Cz^2 + 2Dxy + 2Exz + 2Fyz + 2Gx + 2Hy + 2Iz - l = 0 \]

In this algorithm, we need to check whether a fitted shape is an ellipsoid. In theory, the conditions that ensure a quadratic surface to be an ellipsoid have been well investigated and explicitly stated in analytic geometry textbooks. An ellipsoid can be degenerated into other kinds of elliptic quadrics, such as an elliptic paraboloid. Therefore, a proper constraint must be added. Li and Griffiths gave the following definitions Li and Griffiths (2004).

\[ i = A + B + C \]
\[ j = AB + BC + AC - F^2 - E^2 - D^2 \]

However \(4j - i^2 > 0\) is just a sufficient condition to guarantee that an equation of second degree in three variables represent an ellipsoid, but it is not necessary. In this paper, we assume that these conditions are satisfied.

The algebraic method is a linear problem. It is solving the problem directly and easily. The fitting ellipsoid to a given the data set \((x, y, z)_i, i=1,2,\ldots,n\), is obtained by the solution in the LS sense of in the following:

\[ \Omega v = l \]

Where

\[ \Omega_{xxu} = \text{Design matrix} \]
\[ v_u = [ A \ B \ C \ D \ E \ F \ G \ H \ I ]^T \text{unknown conic parameters} \]
\[ l_n = [1 \ 1 \ 1 \ldots 1]^T \text{ unit vector : right side vector} \]
\[ \text{i}
\text{th row of the } nx9 \text{ matrix } \Omega \]

\[ [x_i^2 \ y_i^2 \ z_i^2 \ 2x_iy_i \ 2x_iz_i \ 2y_iz_i \ 2x_i \ 2y_i \ 2z_i] \]

It is solved easily in the LS sense as below

\[ v = (\Omega^T\Omega)^{-1}\Omega^Tl \]

(12)

or it is solved easily by MATLAB as below

\[ v = [ x^2 \ y^2 \ z^2 \ 2x.*y \ 2x.*z \ 2y.*z \ 2.*x \ 2.*y \ 2.*z] \cdot \text{ones(n)} \]

(13)

If there are differences in weights or correlations between given data points, P weight matrix is added in the solution, and then

\[ v = (\Omega^T \ P \ \Omega)^{-1}\Omega^TPl \]

(14)

\[ P = K_{ll}^{-1}K_{ll}; nxn \text{ variance-covariance matrix of data points} \]

Residual (or correction) vector is computed as below

\[ \varepsilon = [ x^2 \ y^2 \ z^2 \ 2x.y \ 2x.z \ 2y.z ] \cdot v - \text{ones(n)} \]

(15)
LS optimization give us $\|e\| = \text{min}$.

Algebraic methods all have indisputable advantages of solving linear LS problems. The methods for this are well known and fast. However, it is intuitively unclear what it is we are minimizing geometrically in (6) is often referred to as the “algebraic distance” to be minimized Ray and Srivastava (2008). A geometric interpretation given by Bookstein (1979) clearly demonstrates that algebraic methods neglect points far from the center.

5. FITTING OF ELLIPSOID USING ORTHOGONAL DISTANCES

To overcome the problems with the algebraic distances, it is natural to replace them by the orthogonal distances which are invariant to transformations in Euclidean space and which do not exhibit the high curvature bias. An ellipsoid of best fit in the LS sense to the given data points can be found by minimizing the sum of the squares of the geometric distances from the data to the ellipsoid. The geometric distance is defined to be the distance between a data point and its closest point on the ellipsoid.

Determining best fit ellipsoid is a nonlinear least squares problem which in principle can be solved by using the Levenberg-Marquardt (LM) algorithm. Generally, non-linear least squares is a complicated issue. It is very difficult to develop methods which can find the global minimizer with certainty in this situation. When a local minimizer has been discovered, we do not know whether it is a global minimizer or one of the local minimizer Zisserman (2013).

There are a variety of nonlinear optimization techniques. Such as Newton, Gauss-Newton, Gradient Descent, Levenberg-Marquardt approximation etc. However, these fitting techniques involve a highly nonlinear optimization procedure, which often stops at a local minimum and cannot guarantee an optimal solution Li and Griffiths (2004).

Away from the minimum, in regions of negative curvature, the Gauss-Newton approximation is not very good. In such regions, a simple steepest-descent step is probably the best plan. The Levenberg-Marquardt method is a mechanism for varying between steepest-descent and Gauss-Newton steps depending on how good the $H_{\text{GN}}$ approximation is locally.

The Levenberg-Marquardt method uses the modified Hessian

$$H(x, \lambda) = H_{\text{GN}} + \lambda \cdot I$$

where $I$ is the identity matrix.

- When $\lambda$ is small, $H$ approximates the Gauss-Newton Hessian.
- When $\lambda$ is large, $H$ is close to the identity, causing steepest-descent steps to be taken.

This algorithm does not require explicit line searches. More iterations than Gauss-Newton, but, no line search required, and more frequently converge suppose that we have a unknown parameter set

$$v = [ A \quad B \quad C \quad D \quad E \quad F \quad G \quad H \quad I ]^T$$

are unknown conic parameters. The general conic equation for an ellipsoid is given as (6).

We will reach the solution by establishing relationships between variations in the conical coefficients and the orthogonal distances.

The initial parameters were derived from the algebraic fitting ellipsoid.

$$(x_i^*, y_i^*, z_i^*) : \text{Projection coordinates (onto ellipsoid) of given } P_i \text{ data points}$$

$$K = [ x_i^* \quad y_i^* \quad z_i^* \quad I ]$$

$$du = 2.K.[A \quad D \quad E \quad G]^T : \frac{\partial}{\partial x} \text{ partial derivative with respect to } x- \text{ coor.}(16)$$
\[ dv = 2K \begin{bmatrix} D & B & F \end{bmatrix} H^T : \partial / \partial y \text{ partial derivative with respect to y-coor.}(17) \]
\[ dv = 2K \begin{bmatrix} E & F & C \end{bmatrix} I^T : \partial / \partial z \text{ partial derivative with respect to z-coor.}(18) \]
\[ e = (dv^2 + dv^2 + dv^2)^{\frac{1}{2}} \]

\( \text{ith row of the } n \times 10 \text{ matrix } J \) (jakobien matrix)
\[ (I/e) \begin{bmatrix} x_i^2 & y_i^2 & z_i^2 & 2x_iy_i & 2x_i z_i & 2y_i z_i & 2x_i & 2y_i & 2z_i & -1 \end{bmatrix} \]

\( \text{ith row of the right side vector } h_{nx1} \)
\[ h_i = \text{sign}(z_i - z) \cdot \text{sign}(z_i) \cdot \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} \]

We obtained the below linearized equation
\[ J_{nx10} dv_{nx1} = h_{nx1} \]
\[ dv = \begin{bmatrix} dA & dB & dCdDdEdFdGdHeHdI \end{bmatrix}^T \]

The fitted orthogonal ellipsoid is obtained by the solution in the LS sense with the L-M algorithm.

**5.1 The Levenberg-Marquardt Algorithm**

1- Solve algebraic methods and find initial values for \( v \)
\[ \text{set } \lambda = 1 \text{ (say)} \]

2- Compute \( J \)-jakobien matrix and \( h_i \) orthogonal distances from all given data points
\[ \text{minh} = h^T h \]

3- Solve \( (J^T J + \lambda I) dv = J^T h \)
\[ v = v + dv \text{, new conic parameter} \]

Find again \( h_i \) orthogonal distances from all given data points
\[ \text{newh} = h^T h \]

4- ifnewh < minh \% yes there is improvement, reduce \( \lambda \)
\[ \text{minh} = \text{newh}; \lambda = \lambda / 2 \]
\[ \text{goto 3} \]

else \% no improvement, increase \( \lambda \)
\[ \lambda = 2 \cdot \lambda \]
\[ \text{goto 3} \]
end

**5.2. Finding Orthogonal Distances from the Ellipsoid**
In this paper, we present techniques for ellipsoid fitting which are based on minimizing the sum of the squares of the geometric distances between the data and the ellipsoid. The most time-consuming part is the computation of the orthogonal distances between each point and the ellipsoid. Our aim to find the orthogonal distances from a shifted-oriented ellipsoid see Figure 2. For detailed information on this subject refer to Bektas (2014).

6. NUMERICAL EXAMPLE

For numerical applications 12 triplets \((x,y,z)\) cartesian coordinates were produced. Here data points coordinates,
\[
\begin{align*}
x: & \quad [7 \quad 7 \quad 9 \quad 9 \quad 11 \quad 11 \quad 8 \quad 8 \quad 10 \quad 10 \quad 12 \quad 12] \\
y: & \quad [22 \quad 19 \quad 23 \quad 19 \quad 24 \quad 20 \quad 21 \quad 17 \quad 22 \quad 18 \quad 23 \quad 19] \\
z: & \quad [31 \quad 28 \quad 31 \quad 27 \quad 29 \quad 26 \quad 32 \quad 29 \quad 32 \quad 28 \quad 31 \quad 28]
\end{align*}
\]
This problem is also solved by Least Absolute Values Method (\(l_1\)-norm) and the following results were obtained.

The conical coefficients in the Least Squares Method is,
\[
v = \begin{bmatrix}
-0.0006 & -0.0008 & -0.0010 & 0.0005 & -0.0005 & 0.0003 & 0.0092 & 0.0050 & 0.0278 \\
\end{bmatrix}
\]

The conical coefficients in the Least Absolute Values Method is,
\[
v = \begin{bmatrix}
-0.0071 & -0.0084 & -0.0096 & 0.0047 & -0.0040 & 0.0023 & 0.0880 & 0.0061 & 0.0271 \\
\end{bmatrix}
\]

We show both the algebraic and orthogonal fitting results are as shown in Table-1.

<table>
<thead>
<tr>
<th>Algebratic fitting</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Center of Coordinates</td>
<td>Semi-axes</td>
<td>Rotational angles (degree)</td>
<td>RSS*</td>
<td></td>
</tr>
<tr>
<td>(x_0) (y_0) (z_0)</td>
<td>(a_x) (a_y) (b)</td>
<td>(\varepsilon) (\psi) (\omega)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.384</td>
<td>20.965</td>
<td>29.007</td>
<td>7.468</td>
<td>3.164</td>
</tr>
</tbody>
</table>

Orthogonal fitting (after four iterations)

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>11.785</td>
<td>21.694</td>
<td>28.492</td>
<td>11.188</td>
<td>3.346</td>
<td>1.789</td>
<td>47.20</td>
</tr>
</tbody>
</table>

*RSS: The residual sum of squares of the orthogonal distances

7. DISCUSSION

Orthogonal least-squares has a much sounder basis, but is usually difficult to implement. Why are algebraic distances usually not satisfactory? The big advantage of use of algebraic distances is the gain in computational efficiency, because closed-form solutions can usually be obtained. In general, however, the results are not satisfactory. The function to minimize is usually not invariant under Euclidean transformations. For example, the function with normalization \(K' = -1\) in (5) is not invariant with respect to translations. This is a feature we dislike, because we usually do not know in practice where the best coordinate system to represent the data is. A point may contribute differently to the parameter estimation depending on its position on the conic. If a priori all points are corrupted by the same amount...
of noise, it is desirable for them to contribute the same way Zhang (1997). More importantly, algebraic methods have an inherent curvature bias – data corrupted by the same amount of noise will misfit unequally at different curvatures Ray and Srivastava (2008). Our experience tells us that if the coordinates of given points consists of a large number this will cause bad condition. Therefore, before fitting, you must shift the given coordinates to the center of gravity, after fitting operation the coordinates of ellipsoid’s center must be shifted back to the previous position.

8. CONCLUSION

In this paper, we studied on the orthogonal fitting ellipsoid. The problem of fitting ellipsoid is encountered frequently in the image processing, face recognition, computer games, geodesy-determining more realistic Earth ellipsoid etc. The paper has presented a new method of orthogonal fitting ellipsoid. The new method relies on solving an over determined system of nonlinear equations with the use of the L-M method. In conclusion, the presented method may be considered as fast, accurate and reliable and may be successfully used in other areas. The presented orthogonal fitting algorithm can be applied easily to biaxial ellipsoid, sphere and also other surfaces such as paraboloid, hyperboloid, etc.

REFERENCES


Feltens, J. Vector method to compute the Cartesian (X, Y, Z) to geodetic (φ, λ, h) transformation on a triaxial ellipsoid”, Journal of Geod. 83:129–137, 2009.


Zisserman, A. C25 Optimization, 8 Lectures, Hilary Term 2013, 2 Tutorial Sheets, Lectures 3-6 (BK), 2013.