Total Least Square Error Computation in E2: A New Simple, Fast and Robust Algorithm

Vaclav Skala
Department of Computer Science and Engineering
University of West Bohemia
Univerzitni 8, CZ 30614 Plzen, Czech Republic
http://www.VaclavSkala.eu

ABSTRACT

Many problems, not only in signal processing, image processing, digital imaging, computer vision and visualization, lead to the Least Square Error (LSE) problem or Total (Orthogonal) Least Square Error (TLSE) problem computation. Mostly the LSE is used due to its simplicity for problems leading to $f(x,y)=h$, resp. $f(x,y,z)=h$, i.e. to dependences expressible as an explicit function computing "vertical" distances. There are many problems for which the LSE is not convenient and the TSLE is to be used. Those problems usually lead to $F(\boldsymbol{x})=0$, i.e. to dependences expressible as an implicit function computing "orthogonal" distances. Unfortunately, the TLSE is computationally much more expensive.

This paper presents a new, simple, robust and fast algorithm for the total least square error computation in E^2.

CCS Concepts

• **Mathematics of computing → Mathematical analysis**
• **Computing methodologies→Computer vision 3D imaging.**

Keywords

Total least square error; digital imaging; image processing; visualization; computer vision.

1. INTRODUCTION

Wide range of applications is based on approximation of acquired data in E^2 or E^3 space and mostly the Least Square Error minimization is used, known also as a linear or polynomial regression. The regression methods have been heavily explored especially with statistically oriented problems. They are used across many engineering fields dealing with acquired data processing. Several studies have been published and they can be classified as follows:

- "standard" Least Square Error (LSE) methods fitting data to a function $f(\boldsymbol{x})=h$, where $\boldsymbol{x}$ is an independent variable and h is a measured or given value
- "orthogonal" Total Least Square Error (TLSE) methods fitting data to a function $F(\boldsymbol{x})=0$, i.e. fitting data to some $(n-1)$-dimensional entity in this n-dimensional space, e.g. a line in the E^2 space or a plane in the E^3 space [1][6][8][18][19].
- "orthogonally mapping" Total Least Square Error (MTLSE) methods for fitting data to a given entity in a subspace of the given space. However, this problem is much more complicated. As an example we can consider data given in E^n and we need to find an optimal line in E^3, i.e. one dimensional entity, in this n-dimensional space fitting optimally the given data. Typical problem: Find a line in the E^3 space which has the minimum orthogonal distance from the given points in this E^3 space. This approach can be used in optimal parameters estimation, etc. This algorithm is quite complex and solution can be found in [16].

It should be noted that all methods above do have one significant drawback as values are taken in a squared value. This results to an artifact that small values do not have relevant influence to the final entity as the high values. Some methods are trying to overcome this by setting weights to each measured data [3]. It should be noted that the TLSE was originally derived by Pearson [15](1901). Deep comprehensive analysis can be found in [8][12][18]. Differences between the LSE and TLSE methods approaches are significant, see Figure 1.

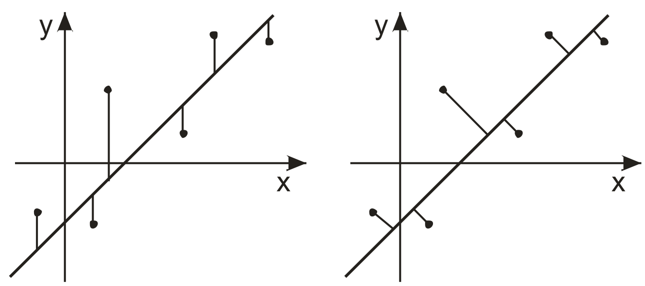

Figure1.a LSE method Figure 1.b TLSE method

In the following we will shortly introduce the Least Square Error method which measures distances "horizontally", than the Total Least Square Error method which measures distances "orthogonally".

2. LEAST SQUARE ERROR

In the vast majority the Least Square Error (LSE) methods measuring vertical distances are used. This approach is acceptable in the case of explicit functional dependences $f(x,y)=h$, resp. $f(x,y,z)=h$. However, it should be noted that a user should keep in a mind, that smaller differences than 1.0, will have significantly smaller weight than higher differences than 1.0 as the differences are taken in a square resulting to dependences in scaling of data approximated, i.e. the result will depend on

Skala,V.: Total Least Square Error Computation in E2: A New Simple, Fast and Robust Algorithm,
CGI 2016 Proceedings, ACM, pp.1-4, doi>10.1145/2949035.2949036, Greece, 2016

physical units used, etc. The main advantage of the LSE method is that it is simple for fitting polynomial curves and it is easy to implement. The standard LSE method leads to over determined system of linear equations. This approach is also known as polynomial regression.

Let us consider a data set $\Omega = \{\langle x_i, y_i, f_i \rangle\}_{i=1}^{n}$, i.e. data set containing for x_i and y_i measured functional value f_i and we want to find parameters $\boldsymbol{a} = [a, b, c, d]^T$ for optimal fitting function, as an example:

$$f(x, y, \boldsymbol{a}) = a + bx + cy + dxy$$

by minimizing the vertical squared distance D, i.e.:

$$D = \min_{a,b,c,d} \sum_{i=1}^{n} \left(f_i - f(x_i, y_i, \boldsymbol{a})\right)^2$$

Conditions for an extreme are given as a vector equation:

$$\frac{\partial D}{\partial \boldsymbol{a}} = \sum_{i=1}^{n} (f_i - (a + bx_i + cy_i + dx_i y_i)) \frac{\partial f(x, y, \boldsymbol{a})}{\partial \boldsymbol{a}} = \boldsymbol{0}$$

Rewriting this vector condition, we obtain conditions:

$$\frac{\partial D}{\partial a} = \sum_{i=1}^{n} (f_i - (a + bx_i + cy_i + dx_i y_i)) = 0$$

$$\frac{\partial D}{\partial b} = \sum_{i=1}^{n} (f_i - (a + bx_i + cy_i + dx_i y_i)) x_i = 0$$

$$\frac{\partial D}{\partial c} = \sum_{i=1}^{n} (f_i - (a + bx_i + cy_i + dx_i y_i)) y_i = 0$$

$$\frac{\partial D}{\partial d} = \sum_{i=1}^{n} (f_i - (a + bx_i + cy_i + dx_i y_i))\, x_i y_i = 0$$

Rewriting the first condition above we get a simple condition:

$$\sum_{i=1}^{n} f_i - na + b \sum_{i=1}^{n} x_i + c \sum_{i=1}^{n} y_i + d \sum_{i=1}^{n} x_i y_i = 0$$

as $\sum_{i=1}^{n} a = na$. Similarly we obtain other conditions.

All those simplified conditions can be rewritten in a matrix form as $\boldsymbol{Ax} = \boldsymbol{b}$. The selection of bilinear form was used to show the LSE method application to a non-linear case, if the case of linear function, i.e. $f(x, y, \boldsymbol{a}) = a + bx + cy$, the 4th row and column of the matrix $\boldsymbol{A}$ is to be removed.

Several methods for LSE have been derived [4][5][10], however those methods are sensitive to the vector $\boldsymbol{a}$ orientation and not robust in general as a value of $\sum_{i=1}^{n} x_i^2 y_i^2$ might be too high in comparison with the value n which has an influence to robustness of numerical solution. Also the LSE methods are sensitive to a rotation as they measure vertical distances. Rotational and translation invariances are fundamental requirements not only in geometrically oriented applications.

3. TOTAL LEAST SQUARE ERROR

The Total (Orthogonal) Least Square (TSLE) method takes another approach as measures distances orthogonally and approximation by a line or plane is used nearly exclusively. One significant property of the TLSE method is its rotational and translational invariance [18][19][19]. This approach leads to an approximation by an implicit function to $F(x, y) = 0$ in the E^2 case, resp. $F(x, y, z) = 0$, in the E^3 case, i.e. to dependences expressible as an implicit function.

There are several approaches how to solve TLSE problem and comprehensive analysis is given in [8]. Many algorithms are based on Singular Value Decomposition (SVD) or on a "simple" solution" based on the explicit line representation [11] in the form $y = bx + a$. This formulation leads to a simple formula for calculation of the a, b coefficients. However, it is not robust and it is sensitive to a rotation. Also when a line is close to a vertical one, there is a high numerical imprecision and an overflow can appear as well, etc. If TLSE method is to be used many times, it is reasonable to consider robust and fast method specialized for the E^2 case. In the E^2 case and the linear case a linear function $F(\boldsymbol{x}) = ax + by + c = 0$ is used, the orthogonal distance d of the given $\boldsymbol{x}$ point and the line p is determined as:

$$d = \frac{|ax + by + c|}{a^2 + b^2}$$

where: $\boldsymbol{x} = [x, y]^T$ is the given point and a line p is given as $ax + by + c = 0$. The computational problem is determination coefficients a, b, c of a line $p \in E^2$.

In image processing, signal processing, digital imaging and computer graphics specialized algorithms should be used in the E^2 case. Such a solution for the E^2 case was published in [2] which is based on a line representation in the polar coordinates. Some specialized algorithms for a circle, resp. ellipse fitting were developed recently as well.The algorithm fully described in [2] is based on polar representation and leads to a formula which is stable. The derivation of the algorithm is not simple and uses goniometric functions, i.e. $\sin(\theta)$ and $\cos(\theta)$. A special case for perfectly circular data is to be solved. The algorithm [2] is not extensible to the E^3 case. In the following a new approach to TLSE computation will be described with experimental verification of the proposed method.

4. PROPOSED ALGORITHM

Fundamental requirement for any algorithm is its robustness. It should be fast and simple to implement as well. The proposed the TLSE algorithm is based on a squared orthogonal distance computation. As the TSLE method has to be translationally and rotationally invariant, the centroid of the given point set is to be $\boldsymbol{x}_0 = \boldsymbol{0}$, this was shown also in [2]. As it is not a general case, the first step is a data set transformation:

$$\boldsymbol{x_i} = \boldsymbol{x_i} - \boldsymbol{x_0} \qquad\qquad \boldsymbol{x_0} = \frac{\boldsymbol{1}}{\boldsymbol{n}} \sum_{i=1}^{n} \boldsymbol{x_i}$$

where: n is a number of the given points, $\boldsymbol{x}_i = [x_i, y_i]^T$ are the given points, $i = 1, \dots, n$. This step has two consequences, the line p passes the origin of the coordinate system and therefore the c coefficient of the line p is set $c = 0$ by definition, now.

There is a seemingly a simple formulation of the TLSE problem using optimization and Lagrange multipliers, i.e.

$$\min_{a,b,\lambda} D(a, b, \lambda) = \min_{a,b,\lambda} \{(ax_i + by_i)^2 + \lambda g(a, b)\} \;\&\, g(a, b) = 0$$

where: $g(a, b) = a^2 + b^2 - 1$. Unfortunately, this approach does not lead to a simple solution.

The proposed algorithm is based on direct minimization of a distance D given as:

$$D(a, b) = \sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} \frac{(ax_i + by_i)^2}{a^2 + b^2}$$

For a minimum the following conditions must be fulfilled

$$\frac{\partial D(a,b)}{\partial a} = 0 \qquad \frac{\partial D(a,b)}{\partial b} = 0$$

Now, we can express both conditions analytically as follows

$$\frac{\partial D(a,b)}{\partial a} = 0 = \frac{\sum_{i=1}^{n}[2(ax_i + by_i)x_i(a^2+b^2)] - \sum_{i=1}^{n}[(ax_i+by_i)^2 2a]}{(a^2+b^2)^2}$$

Multiplying by $(a^2+b^2)^2 \neq 0$ and by $1/2$ we get

$$\frac{\partial D(a,b)}{\partial a} = \sum_{i=1}^{n}[(ax_i+by_i)x_i(a^2+b^2)] - \sum_{i=1}^{n}[(ax_i+by_i)^2 a]$$

$$= \sum_{i=1}^{n}[(ax_i+by_i)x_i(a^2+b^2) - a(ax_i+by_i)^2] = 0$$

$$\sum_{i=1}^{n}\left[\left(ax_i^2 + bx_iy_i\right)(a^2+b^2) - a\left(a^2x_i^2 + b^2y_i^2 + 2abx_iy_i\right)\right] = 0$$

Using algebraic operations we get:

$$\frac{\partial D(a,b)}{\partial a} = \sum_{i=1}^{n}\left(a^3x_i^2 + a^2b\,x_iy_i + ab^2x_i^2 + b^3y_i^2 - a^3x_i^2 - a^2b\,y_i^2 - 2a^2b\,x_iy_i\right) = 0$$

and by simplifying we obtain:

$$\frac{\partial D(a,b)}{\partial a} = \sum_{i=1}^{n}\left[ab^2\left(x_i^2 - y_i^2\right) - a^2b\,x_iy_i + b^3y_i^2\right] = 0$$

We can divide the equation by $b \neq 0$ and then

$$\frac{\partial D(a,b)}{\partial a} = \sum_{i=1}^{n}\left[ab\left(x_i^2 - y_i^2\right) + x_iy_i(b^2-a^2)\right] = 0$$

and it can be rewritten as

$$ab\sum_{i=1}^{n}\left(x_i^2 - y_i^2\right) + (b^2-a^2)\sum_{i=1}^{n}x_iy_i = 0$$

The second condition for an extreme is given as:

$$\frac{\partial D(a,b)}{\partial b} = 0 = \frac{\sum_{i=1}^{n}[2(ax_i + by_i)y_i(a^2+b^2)] - \sum_{i=1}^{n}[(ax_i+by_i)^2 2b]}{(a^2+b^2)^2}$$

Multiplying by $(a^2+b^2)^2/2 \neq 0$ we get

$$\frac{\partial D(a,b)}{\partial b} = \sum_{i=1}^{n}[(ax_i+by_i)y_i(a^2+b^2)] - \sum_{i=1}^{n}[(ax_i+by_i)^2 b]$$

$$= \sum_{i=1}^{n}[(ax_i+by_i)y_i(a^2+b^2) - b(ax_i+by_i)^2]$$

$$= \sum_{i=1}^{n}\left[\left(ax_iy_ix_i^2 + by_i^2\right)(a^2+b^2) - b\left(a^2x_i^2 + b^2y_i^2 + 2abx_iy_i\right)\right] = 0$$

Using algebraic operations we get:

$$\frac{\partial D(a,b)}{\partial b} = \sum_{i=1}^{n}\left(a^3x_iy_i + a^2by_i^2 + ab^2x_iy_i + b^3y_i^2 - a^2bx_i^2 - b^3\,y_i^2 - 2ab^2x_iy_i\right) = 0$$

and by simplifying we obtain:

$$\frac{\partial D(a,b)}{\partial b} = \sum_{i=1}^{n}\left[a^3\,x_iy_i + ab^2\left(y_i^2 - x_i^2\right) - ab^2x_iy_i\right] = 0$$

We can divide the equation by $a \neq 0$ and then

$$\frac{\partial D(a,b)}{\partial b} = \sum_{i=1}^{n}\left[ab\left(x_i^2 - y_i^2\right) + x_iy_i(b^2-a^2)\right] = 0$$

and it can be rewritten as:

$$ab\sum_{i=1}^{n}\left(x_i^2 - y_i^2\right) + (b^2-a^2)\sum_{i=1}^{n}x_iy_i = 0$$

Now, we have got two equations from those two conditions for an extreme:

$$\frac{\partial D(a,b)}{\partial a} = ab\sum_{i=1}^{n}\left(x_i^2 - y_i^2\right) + (b^2-a^2)\sum_{i=1}^{n}x_iy_i = 0$$

$$\frac{\partial D(a,b)}{\partial b} = ab\sum_{i=1}^{n}\left(x_i^2 - y_i^2\right) + (b^2-a^2)\sum_{i=1}^{n}x_iy_i = 0$$

It can be seen that both equations above are equivalent and actually we have got just one equation:

$$ab\sum_{i=1}^{n}\left(x_i^2 - y_i^2\right) + (b^2-a^2)\sum_{i=1}^{n}x_iy_i = 0$$

which can be rewritten using substitutions as:

$$\xi = \sum_{i=1}^{n}\left(x_i^2 - y_i^2\right) \qquad \eta = \sum_{i=1}^{n}x_iy_i \qquad ab\xi + (b^2-a^2)\eta = 0$$

Now, we need to determine values a, b. As we keep the normalization condition for coefficients a, b during the extreme conditions, if $\sum_{i=1}^{n}x_i^2 > \sum_{i=1}^{n}y_i^2$ we can select the value a, e.g. $a = 1$, and solve the equation for b or vice versa. This leads to a quadratic equation:

$$b\xi + (b^2-1)\eta = 0 \qquad \text{i.e.} \qquad \eta b^2 + \xi b - \eta = 0$$

and therefore

$$b_1 = \frac{-\xi + \sqrt{\xi^2 - 4\eta}}{2\eta} \qquad b_2 = \frac{-\xi - \sqrt{\xi^2 - 4\eta}}{2\eta}$$

The minimum distance is given by the b_2 value and

$$b = b_2 \qquad a = 1$$

and the a, b values are of general values.

Now, the computed line $p: ax + by = 0$, which is represented by the vector $\boldsymbol{p} = [a, b: 0]^T$, passes the origin of the coordinated system is to be "moved" back to the original coordinate system of the original data set using the standard geometric transformation represented by a matrix $\boldsymbol{T}$ [17], i.e.:

$$\begin{bmatrix} a' \\ b' \\ c' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -x_0 & -y_0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} \qquad \boldsymbol{p}' = \boldsymbol{T}\boldsymbol{p}$$

Now, the line $\boldsymbol{p}'$ represents the line which optimally fits data in the sense of the TLSE method.

The formula is simple, easy and robust. However, it should be noted, that the proposed method above is not directly extensible to the E^3 case.

5. EXPERIMENTAL VERIFICATION

The proposed TLSE algorithm for the E^2 case was tested on datasets with known properties in order to verify the correctness and robustness, also for randomly generated data using MATLAB system. The experiments proved expected properties and the correctness of the proposed algorithm and its expected properties, Figure 2.

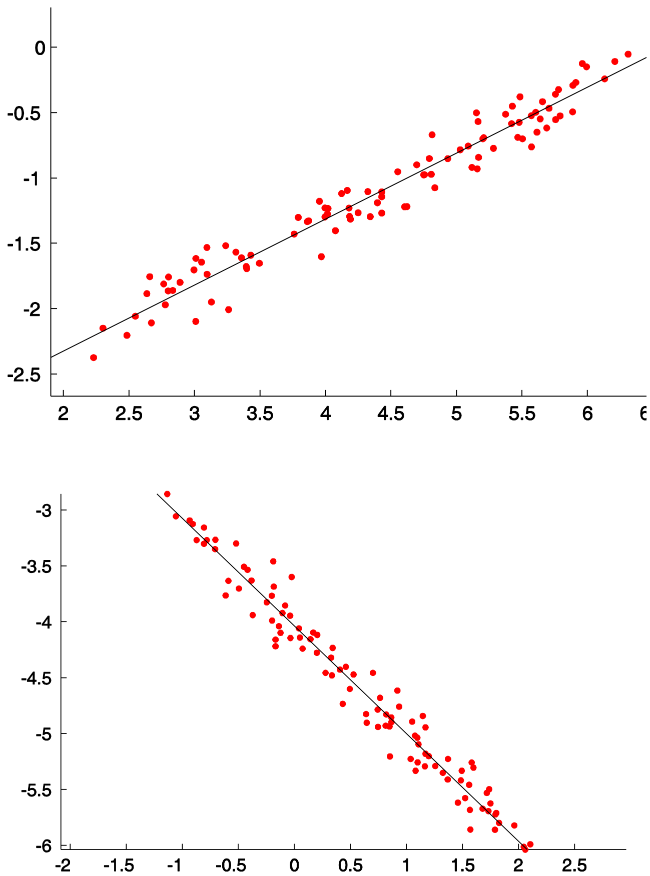

Figure 2 Examples of the proposed TLSE application

In the case of large data sets, it is recommended for the sake of precision to compute the expression $\sqrt{\xi^2 - 4\eta}$ in a double precision than the rest of the formula, according to standard numerical mathematics recommendations.

6. CONCLUSION

In this paper a new simple, fast and robust method for approximation using Total Least Square Error method in E^2 has been presented. The method was experimentally verified using the MATLAB system. However, it should be noted, that the presented method is limited to the E^2 case, only. The proposed TLSE method is significantly simpler than the "standard" Least Square Error method.

In the E^3 case, the TLSE problem is more complicated and no simple analogous formula has been derived, yet.

7. ACKNOWLEDGMENTS

The author would like to thank to colleagues at the University of West Bohemia in Plzen for fruitful discussions and to anonymous reviewers for their comments and hints which helped to improve the manuscript significantly. Special thanks belong to Michal Šmolík for his independent algorithm verification and graphs generation in MATLAB. This research was supported by the MSMT CR - project No. LH12181.

8. REFERENCES

[1] Abatzoglou,T., Mendel,J. 1987. Constrained total least squares, in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP'87), Vol. 12, 1485–1488.

[2] Alciatore,D., Miranda,R. 1995. The Best Least-Square Line Fit, Graphics Gems V (Ed.Paeth,A.W.), 91-97, Academic Press.

[3] Amiri-Simkooei,A.R., Jazaeri,S. 2012. Weighted total least squares formulated by standard least squares theory, Journal of Geodetic Science, 2 (2): 113-124.

[4] Charpa,S., Canale,R. 1988: Numerical methods for Engineers, McGraw-Hill.

[5] Chatfield,C. 1970. Statistics for Technology, Penguin Book.

[6] de Groen,P. 1996 An introduction to total least squares, Nieuw Archief voor Wiskunde,Vierde serie, deel 14, 237–253

[7] DeGroat,R.D., Dowling,E.M. 1993 The data least squares problem and channel equalization. IEEE Trans. Signal Processing, Vol. 41, No. 1, 407–411.

[8] Golub,G.H., Van Loan,C.F. 1980. An analysis of the total least squares problem. SIAM J. on Numer. Anal., 17, 883–893.

[9] Jo,S., Kim,S.W. 2005. Consistent normalized least mean square filtering with noisy data matrix. IEEE Trans. Signal Processing, Vol. 53, No. 6, 2112–2123.

[10] Kryszig,V. 1983. Advanced Engineering Mathematics, John Wiley & Sons.

[11] Lee,S.L. 1994. A note on the Total Least Square Fit to coplanar points, Tech.Rep. ORNL-TM-12852, Oak Ridge National Laboratory.

[12] Nievergelt,Y. 1994. Total Least Squares: State of the Art regression in numerical mathematics, SIAM Review, Vol.36, 258-264

[13] Nixon,M.S., Aguado,A.S. 2012. Feature Extraction & Image processing for Computer Vision, Academic Press.

[14] Markowsky,I., VanHueffel,S. 2007. Overview of total least square methods, Signal Processing, 87 (10), 2283-2302.

[15] Pearson,K. 1901. On line and planes of closest fit to system of points in space, Phil.Mag., Vol.2, 559-572

[16] Skala,V. 2015. New Formulation for Total Least Square Error Method in d-dimensional Space with Mapping to a Parametric Line, ICNAAM 2015 conference, accepted for publication AIP Conf. Proceedings

[17] Skala,V. 2012. Projective Geometry and Duality for Graphics, Games and Visualization - Course SIGGRAPH Asia 2012, Singapore.

[18] Van Huffel,S., Vandewalle,J 1991. The Total Least Squares Problems: Computational Aspects and Analysis. SIAM Publications, Philadelphia PA.

[19] Van Huffel,S., Lemmerling,P. 2002. Total Least Squares and Errors-in-Variables Modeling: Analysis, Algorithms and Applications. Dordrecht, The Netherlands: Kluwer Academic Publishers

[20] Perpendicular Regression of a Line (download 2015-12-20) http://www.mathpages.com/home/kmath110.htm