

Article

A Polynomial Fitting Problem: The Orthogonal Distances Method

Luis Alberto Cantera-Cantera ^{1,*}, Cristóbal Vargas-Jarillo ², Sergio Isai Palomino-Reséndiz ¹,
Yair Lozano-Hernández ³ and Carlos Manuel Montelongo-Vázquez ¹

¹ Automation and Control Engineering Department, ESIME-IPN, Unidad Profesional Adolfo López Mateos, Zacatenco, Av. Luis Enrique Erro s/n, Mexico City C.P. 07738, Mexico

² Automatic Control Department, CINVESTAV-IPN, Av. Instituto Politecnico Nacional 2508, Col. San Pedro Zacatenco, Mexico City C.P. 07360, Mexico

³ Unidad Profesional Interdisciplinaria de Ingeniería Campus Hidalgo, Instituto Politécnico Nacional, Hidalgo 42162, Mexico

* Correspondence: lcanterac@ipn.mx

Abstract: The classical curve-fitting problem to relate two variables, x and y , deals with polynomials. Generally, this problem is solved by the least squares method (LS), where the minimization function considers the vertical errors from the data points to the fitting curve. Another curve-fitting method is total least squares (TLS), which takes into account errors in both x and y variables. A further method is the orthogonal distances method (OD), which minimizes the sum of the squares of orthogonal distances from the data points to the fitting curve. In this work, we develop the OD method for the polynomial fitting of degree n and compare the TLS and OD methods. The results show that TLS and OD methods are not equivalent in general; however, both methods get the same estimates when a polynomial of degree 1 without an independent coefficient is considered. As examples, we consider the calibration curve-fitting problem of a R-type thermocouple by polynomials of degrees 1 to 4, with and without an independent coefficient, using the LS, TLS and OD methods.

Keywords: polynomial fitting; parameter estimation; orthogonal distances; least squares; total least squares

MSC: 65D10; 65D15



Citation: Cantera-Cantera, L.A.; Vargas-Jarillo, C.; Palomino-Reséndiz, S.I.; Lozano-Hernandez, Y.; Montelongo-Vázquez, C.M. A Polynomial Fitting Problem: The Orthogonal Distances Method. *Mathematics* **2022**, *10*, 4596. <https://doi.org/10.3390/math10234596>

Academic Editors: Juan R. Torregrosa, Alicia Cordero and Francisco I. Chicharro

Received: 25 October 2022

Accepted: 28 November 2022

Published: 4 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The curve-fitting or parameter estimation procedure consists of determining the parameters of a curve to obtain the best possible fit to observations. Depending on the curve, this problem can be classified as linear or nonlinear [1,2]. The most used curves for fitting in biology, physics, statistics and engineering are polynomials $y = a_0 + a_1x + \dots + a_{n-1}x^{n-1} + a_nx^n$ [3–7]. The polynomial fitting is a linear problem and the most popular method used to solve it is least squares (LS) [8–10], which minimizes the sum of the squares of vertical errors from the data points $(x_1, y_1), \dots, (x_m, y_m)$ to the polynomial curve. Another method that can be used to solve the problem is total least squares (TLS), which considers that both x and y are subject to errors; this method was developed mainly by Golub and Van Loan [11] and Van Huffel and Vandewalle [12]; however, the consideration that both variables x and y are subject to errors dates back at least to 1943 in the work of Deming [13]. Another way to solve the polynomial fitting is to consider the minimization of the sum of the squares of orthogonal distances from the data points to the polynomial curve, called in this article as the orthogonal distances (OD) method. Adcock [14] was the first to use the OD method for the straight line fitting; however, it has also been studied and used for the straight line and planes fitting in [15–21].

The remainder of this paper is structured as follows: Section 2 presents a brief description of the LS and TLS methods. The OD method is developed in Section 3 and it is also

compared with the TLS method using the polynomial of degree 1. The approximation of the output voltage corresponding to different temperatures of a R-type thermocouple [22] using polynomials of degrees 1 to 4 by LS, TLS and OD methods is carried out in Section 4. Finally, a discussion and concluding remarks are presented in Sections 5 and 6, respectively.

In general, this article studies the polynomial fitting problem by the OD method, and the main contributions are the following:

- Development of the OD method for polynomial fitting of degree n .
- Polynomial fitting of thermoelectric voltage data from a R-type thermocouple by OD, LS and TLS methods.
- From the numerical experiments studied, the TLS and OD methods are not equivalent in general. However, both methods obtain the same estimates when the polynomial of degree 1 without an independent coefficient is used.

2. Polynomial Fitting by LS and TLS Methods

The polynomial fitting of degree n consists of determining the unknown parameters $a_0, a_1, \dots, a_n$ of the polynomial defined for all real numbers x , such that $f(x) = a_0 + a_1x + \dots + a_{n-1}x^{n-1} + a_nx^n$. A brief description of the LS and TLS solutions for polynomial fitting are presented below.

2.1. The LS Method

The LS method is one of the most widely used minimization criteria for parameter estimation. The first publication of the method was made by A. M. Legendre in 1805; however, C. F. Gauss claimed that he used the method 10 years earlier [16]. The classical LS method seeks to determine the curve $y = f(x)$, which best fits to observations (x_i, y_i) with $i = 1, \dots, m$, minimizing the sum of the squares of vertical errors defined as $r_i = y_i - f(x_i)$; however, there are some variants of the method, such as weighted least squares, generalized least squares and maximum likelihood, among others, and it can be applied to different mathematical models, such as algebraic models, and continuous and discrete dynamic models [23–28] using the off-line and recursive schemes [20,29,30].

Since the article considers the classical LS method under the off-line scheme, the vector of vertical errors for the polynomial of degree n and observations (x_i, y_i) is given by

$$\mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} - \begin{bmatrix} 1 & x_1 & \cdots & x_1^n \\ 1 & x_2 & \cdots & x_2^n \\ \vdots & \vdots & & \vdots \\ 1 & x_m & \cdots & x_m^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

Defining

$$\mathbf{y} := \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}, \mathbf{A} := \begin{bmatrix} 1 & x_1 & \cdots & x_1^n \\ 1 & x_2 & \cdots & x_2^n \\ \vdots & \vdots & & \vdots \\ 1 & x_m & \cdots & x_m^n \end{bmatrix}, \mathbf{p} := \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix},$$

we obtain

$$\mathbf{r} = \mathbf{y} - \mathbf{A}\mathbf{p}, \quad (1)$$

where $\mathbf{A} \in \mathbb{R}^{m \times (n+1)}$, $\mathbf{y} \in \mathbb{R}^m$ and $\mathbf{p} \in \mathbb{R}^{(n+1)}$.

Taking the squared 2-norm of vector (1), we obtain

$$\|\mathbf{r}\|_2^2 = \|\mathbf{y} - \mathbf{A}\mathbf{p}\|_2^2 = \sum_{i=1}^m r_i^2. \quad (2)$$

The LS solution is the vector $\mathbf{p}$ such that it minimizes the functional (2). Since it is a linear minimization problem, the LS solution can be found by elementary calculus [10,31,32]. The $\mathbf{p}$ LS solution which minimizes (2) is unique if matrix $\mathbf{A}$ has independent columns, which means that it has rank $\rho(\mathbf{A}) = n + 1$. Therefore, the $\mathbf{p}$ LS solution satisfies the normal equations $\mathbf{A}^\top \mathbf{A} \mathbf{p} = \mathbf{A}^\top \mathbf{y}$ and is defined by $\mathbf{p} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{y}$.

On the other hand, if matrix $\mathbf{A}$ has dependent columns or is not invertible, the $\mathbf{p}$ LS solution which minimizes (2) is defined by

$$\mathbf{p} = \mathbf{A}^\dagger \mathbf{y}, \quad (3)$$

where $\mathbf{A}^\dagger$ is called the Moore–Penrose pseudoinverse of $\mathbf{A}$, and it is defined by

$$\mathbf{A}^\dagger = V \begin{bmatrix} \Sigma_q^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} U^\top.$$

Equation (3) is called the pseudoinverse solution of the LS method, and the Moore–Penrose pseudoinverse is defined by the matrices of the singular value decomposition (SVD) given by the following theorem. Throughout the text, the rank of any matrix will be denoted by $\rho(\cdot)$.

Theorem 1 ([10]). *Let any matrix $\mathbf{A} \in \mathbb{R}^{m \times (n+1)}$. The SVD of $\mathbf{A}$ is*

$$\mathbf{A} = U \begin{bmatrix} \Sigma_q & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} V^\top,$$

where $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{(n+1) \times (n+1)}$ are unitary matrices and $\Sigma_q = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_q)$. The numbers $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_q > 0$ are called the singular values of $\mathbf{A}$ and $\rho(\mathbf{A}) = q$.

More properties of the Moore–Penrose pseudoinverse can be found in [10,32].

2.2. The TLS Method

This section presents the most important results of the TLS method from [12] for polynomial fitting.

Rewriting the system of linear algebraic equations $\mathbf{A} \mathbf{p} = \mathbf{y}$ as follows

$$[\mathbf{A} \quad \mathbf{y}] \begin{bmatrix} \mathbf{p} \\ -1 \end{bmatrix} = \mathbf{0},$$

since $\mathbf{A} \in \mathbb{R}^{m \times (n+1)}$, $\mathbf{y} \in \mathbb{R}^m$ and $\mathbf{p} \in \mathbb{R}^{(n+1)}$, it can be defined $\mathbf{\Lambda} := [\mathbf{A} \quad \mathbf{y}] \in \mathbb{R}^{m \times (n+2)}$ and $\mathbf{z} := \begin{bmatrix} \mathbf{p} \\ -1 \end{bmatrix} \in \mathbb{R}^{(n+2)}$ to obtain

$$\mathbf{\Lambda} \mathbf{z} = \mathbf{0}. \quad (4)$$

Since Equation (4) represents a homogeneous system of linear algebraic equations, the $\mathbf{z}$ solution is in the null space of $\mathbf{\Lambda}$, which means that the parameters vector $\mathbf{p}$ is also in the null space of $\mathbf{\Lambda}$. If matrix $\mathbf{\Lambda}$ in (4) has rank $\rho(\mathbf{\Lambda}) = n + 2$, the SVD of $\mathbf{\Lambda}$ is given by

$$\mathbf{\Lambda} = U \Sigma V^\top = \sum_{i=1}^{n+2} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top,$$

where the singular values σ_i and singular vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ satisfy

$$\begin{aligned}
\Lambda \mathbf{v}_1 &= \sigma_1 \mathbf{u}_1, \\
\Lambda \mathbf{v}_2 &= \sigma_2 \mathbf{u}_2, \\
&\vdots \\
\Lambda \mathbf{v}_{n+1} &= \sigma_{n+1} \mathbf{u}_{n+1}, \\
\Lambda \mathbf{v}_{n+2} &= \sigma_{n+2} \mathbf{u}_{n+2}.
\end{aligned} \tag{5}$$

From (5), it can be seen that system (4) does not have a vector $\mathbf{v}$ in the null space of Λ , and therefore the TLS method searches for a matrix Λ_k with $\rho(\Lambda_k) = k < n + 2$, such that $\|\Lambda - \Lambda_k\|$ is minimal. Once a minimizing Λ_k is found, any $\mathbf{z}$ vector satisfying $\Lambda_k \mathbf{z} = \mathbf{0}$ is called the TLS solution and $\Delta\Lambda = \Lambda - \Lambda_k$ is the TLS correction.

The minimization is found by the Eckart–Young–Mirsky theorem. This theorem was initially proved by Eckart and Young [33] using the Frobenius norm, and later Mirsky [34] proved it using the 2-norm. In this work, it is used with the 2-norm. The Eckart–Young–Mirsky theorem using the 2-norm is presented below.

Theorem 2 ([12]). *Let the SVD of $\Lambda \in \mathbb{R}^{m \times (n+2)}$ be given by $\Lambda = U\Sigma V^\top$ with $\rho(\Lambda) = q$ and let $k < q$; then, the matrix*

$$\Lambda_k = \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^\top,$$

satisfies

$$\min_{\rho(\Lambda)=k} \|\Lambda - \mathbf{B}\|_2 = \|\Lambda - \Lambda_k\|_2 = \sigma_{k+1}.$$

Applying the Eckart–Young–Mirsky theorem 2 to (4), the best matrix approximation Λ_k to the matrix Λ is given by

$$\Lambda_k = \sum_{i=1}^{n+1} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top, \quad \sigma_1 \geq \cdots \geq \sigma_{n+1} > 0.$$

Notice that Λ_k has rank $\rho(\Lambda_k) = n + 1$, then the singular values σ_i and singular vectors $\mathbf{u}_i, \mathbf{v}_i$ now satisfy

$$\begin{aligned}
\Lambda_k \mathbf{v}_1 &= \sigma_1 \mathbf{u}_1, \\
\Lambda_k \mathbf{v}_2 &= \sigma_2 \mathbf{u}_2, \\
&\vdots \\
\Lambda_k \mathbf{v}_{n+1} &= \sigma_{n+1} \mathbf{u}_{n+1}, \\
\Lambda_k \mathbf{v}_{n+2} &= \mathbf{0}.
\end{aligned} \tag{6}$$

From Equation (6), the vector $\mathbf{v}_{n+2}$ is in the null space of Λ_k , and therefore, a $\mathbf{z}$ TLS solution satisfying $\Lambda_k \mathbf{z} = \mathbf{0}$ is given by

$$\mathbf{z} = \begin{bmatrix} \mathbf{p} \\ -1 \end{bmatrix} = \mathbf{v}_{n+2} = \begin{bmatrix} \mathbf{v}_{1,n+2} \\ \mathbf{v}_{2,n+2} \\ \vdots \\ \mathbf{v}_{n+2,n+2} \end{bmatrix}. \tag{7}$$

The TLS solution will be scaled to obtain -1 in the last row of (7):

$$\mathbf{z} = \begin{bmatrix} \mathbf{p} \\ -1 \end{bmatrix} = -\frac{1}{\mathbf{v}_{n+2,n+2}} \begin{bmatrix} \mathbf{v}_{1,n+2} \\ \mathbf{v}_{2,n+2} \\ \vdots \\ \mathbf{v}_{n+2,n+2} \end{bmatrix},$$

then, if $\mathbf{v}_{n+2,n+2} \neq 0$, the parameters vector $\mathbf{p}$ is given by

$$\mathbf{p} = -\frac{1}{\mathbf{v}_{n+2,n+2}} \begin{bmatrix} \mathbf{v}_{1,n+2} \\ \mathbf{v}_{2,n+2} \\ \vdots \\ \mathbf{v}_{n+1,n+2} \end{bmatrix}. \quad (8)$$

An alternative form to obtain (8) is as follows. Since the $\mathbf{z}$ solution is also an eigenvector of $\mathbf{\Lambda}^\top \mathbf{\Lambda}$, it is associated with the eigenvalue σ_{n+2}^2 . Hence the eigenvalue–eigenvector equation is given by

$$\mathbf{\Lambda}^\top \mathbf{\Lambda} \begin{bmatrix} \mathbf{p} \\ -1 \end{bmatrix} = \sigma_{n+2}^2 \begin{bmatrix} \mathbf{p} \\ -1 \end{bmatrix},$$

substituting $\mathbf{\Lambda} := [\mathbf{A} \ \mathbf{y}]$ we obtain

$$\begin{bmatrix} \mathbf{A}^\top \mathbf{A} & \mathbf{A}^\top \mathbf{y} \\ \mathbf{y}^\top \mathbf{A} & \mathbf{y}^\top \mathbf{y} \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ -1 \end{bmatrix} = \sigma_{n+2}^2 \begin{bmatrix} \mathbf{p} \\ -1 \end{bmatrix}. \quad (9)$$

Selecting the first row of Equation (9)

$$\mathbf{A}^\top \mathbf{A} \mathbf{p} - \mathbf{A}^\top \mathbf{y} = \sigma_{n+2}^2 \mathbf{p},$$

and solving for the parameters vector $\mathbf{p}$, we obtain

$$\mathbf{p} = (\mathbf{A}^\top \mathbf{A} - \sigma_{n+2}^2 \mathbf{I})^{-1} \mathbf{A}^\top \mathbf{y}. \quad (10)$$

The following theorem establishes the uniqueness and existence conditions of the TLS solution.

Theorem 3 ([12]). *Let the singular values of $\mathbf{A}$ be $\hat{\sigma}_1 \geq \hat{\sigma}_2 \geq \dots \geq \hat{\sigma}_{n+1} > 0$, and the singular values of $[\mathbf{A} \ \mathbf{y}]$ be $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{n+2} > 0$. If $\hat{\sigma}_{n+1} > \sigma_{n+2}$ then the TLS problem has a unique solution.*

As we can see, one problem of the TLS solution is the loss of uniqueness due to repeated singular values, and another is the indeterminate form in (8) if $\mathbf{v}_{n+2,n+2} = 0$; those problems are called nongeneric TLS problems, and more details of them can be found in [35–37].

3. Polynomial Fitting by OD Method

The OD method minimizes the sum of the squares of the orthogonal distances from observations $(x_1, y_1), \dots, (x_m, y_m)$ to the polynomial curve $y = f(x)$. Considering Figure 1, where $\mathbf{a}_1$ is the vector of the observation (x_1, y_1) , $\mathbf{c}$ is a vector on the curve $(x, f(x))$, $\mathbf{c}' := \frac{d}{dx} \mathbf{c} = (1, f'(x))$ is the tangent vector at $\mathbf{c}$ and the vector $\mathbf{d}_1$ is defined by

$$\mathbf{d}_1 = \mathbf{a}_1 - \mathbf{c}. \quad (11)$$

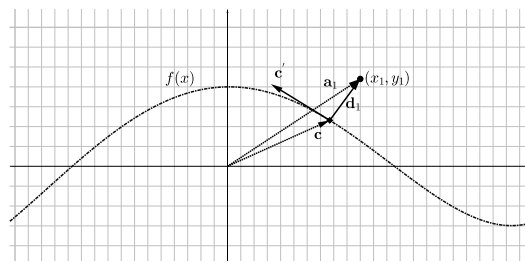


Figure 1. Orthogonal distance from a point to curve.

The vector $\mathbf{d}_1$ in Figure 1 is orthogonal to the curve $f(x)$ at c if the dot product between vectors $\mathbf{d}_1$ and $\mathbf{c}'$ satisfies

$$\mathbf{d}_1 \cdot \mathbf{c}' = 0. \quad (12)$$

Substituting $\mathbf{c}'$ and $\mathbf{d}_1$ in (12), we obtain

$$\mathbf{d}_1 \cdot \mathbf{c}' = x_1 - x + f'(x)(y_1 - f(x)) = 0. \quad (13)$$

Equation (13) is called the orthogonality condition, and its solution determines the point $(x_1^*, f(x_1^*))$ on the curve $f(x)$ where the vectors $\mathbf{d}_1$ and $\mathbf{c}'$ are orthogonal. This point also determines the orthogonal distance vector $\mathbf{d}_1$. The 2-norm of vector $\mathbf{d}_1$ is the orthogonal distance from the observation (x_1, y_1) to the curve $f(x)$ denoted by $\|\mathbf{d}_1\|_2$.

If the polynomial curve considered is $f_1(x) = a_0 + a_1x$, then by the orthogonality condition (13) and the squared 2-norm of vector $\mathbf{d}_i$ (11), the squares of the orthogonal distances for the observations (x_i, y_i) with $i = 1, \dots, m$ are given by

$$\|\mathbf{d}_i\|_2^2 = \frac{r_i^2}{1 + a_1^2},$$

with $r_i = y_i - (a_0 + a_1x_i)$. Defining the distances vector $\mathbf{d} = [\|\mathbf{d}_1\|_2 \quad \|\mathbf{d}_2\|_2 \quad \dots \quad \|\mathbf{d}_m\|_2]^\top$ and taking the squared 2-norm of $\mathbf{d}$, we obtain

$$\|\mathbf{d}\|_2^2 = \mathbf{d} \cdot \mathbf{d} = \frac{1}{1 + a_1^2} \sum_{i=1}^m r_i^2. \quad (14)$$

Equation (14) is the functional of the sum of the squares of the orthogonal distances from observations (x_i, y_i) to the polynomial curve $y = a_0 + a_1x$. The parameters a_0, a_1 that minimize the functional (14) are defined below.

Considering the function

$$f(a_0, a_1) := \frac{1}{1 + a_1^2} \sum_{i=1}^m r_i^2, \quad (15)$$

with $r_i = y_i - (a_0 + a_1x_i)$. Differentiating (15) with respect to a_1 and a_0 and equating to zero

$$\begin{aligned} f_{a_1} := \frac{\partial f}{\partial a_1} &= - \sum_{i=1}^m x_i y_i + a_1 \sum_{i=1}^m (x_i^2 - y_i^2) + a_0 \sum_{i=1}^m x_i \\ &\quad + a_1^2 \sum_{i=1}^m x_i y_i - a_1^2 a_0 \sum_{i=1}^m x_i \\ &\quad + 2a_1 a_0 \sum_{i=1}^m y_i - m a_1 a_0^2 = 0. \end{aligned} \quad (16)$$

$$f_{a_0} := \frac{\partial f}{\partial a_0} = \sum_{i=1}^m y_i - a_1 \sum_{i=1}^m x_i - ma_0 = 0. \quad (17)$$

From Equation (17) we obtain

$$a_0 = \frac{1}{m} \left(\sum_{i=1}^m y_i - a_1 \sum_{i=1}^m x_i \right). \quad (18)$$

Substituting (18) into (16) boils down to

$$\mu_1 a_1^2 + \mu_2 a_1 + \mu_3 = 0, \quad (19)$$

where

$$\begin{aligned} \mu_1 &= \sum_{i=1}^m x_i y_i - \frac{1}{m} \sum_{i=1}^m y_i \sum_{i=1}^m x_i, \\ \mu_2 &= \sum_{i=1}^m (x_i^2 - y_i^2) + \frac{1}{m} \left(\sum_{i=1}^m y_i \right)^2 - \frac{1}{m} \left(\sum_{i=1}^m x_i \right)^2, \\ \mu_3 &= \frac{1}{m} \sum_{i=1}^m y_i \sum_{i=1}^m x_i - \sum_{i=1}^m x_i y_i. \end{aligned}$$

Solving quadratic Equation (19)

$$a_1^{(1,2)} = \frac{-\mu_2 \pm \sqrt{\mu_2^2 - 4\mu_1\mu_3}}{2\mu_1} \quad (20)$$

and substituting (20) into (18), we obtain

$$a_0^{(1,2)} = \frac{1}{m} \left(\sum_{i=1}^m y_i - a_1^{(1,2)} \sum_{i=1}^m x_i \right). \quad (21)$$

Since the sets of values $(a_1^{(1)}, a_0^{(1)})$ and $(a_1^{(2)}, a_0^{(2)})$ are the critical points of function (15), the second derivative test is used to determine the minimum.

Definition 1. Let $f(a_0, a_1)$ be a differentiable function with respect to a_0 and a_1 . Let $H_f(a_0^*, a_1^*) = f_{a_1 a_1}(a_0^*, a_1^*) f_{a_0 a_0}(a_0^*, a_1^*) - [f_{a_0 a_1}(a_0^*, a_1^*)]^2$. If (a^*, b^*) is a critical point of f , that is, $f_{a_0}(a_0^*, a_1^*) = 0$ and $f_{a_1}(a_0^*, a_1^*) = 0$. Then

- If $H_f(a_0^*, a_1^*) > 0$ and $f_{a_1 a_1}(a^*, b^*) > 0$, then $f(a_0^*, a_1^*)$ is a local minimum value of f .
- If $H_f(a_0^*, a_1^*) > 0$ and $f_{a_1 a_1}(a^*, b^*) < 0$, then $f(a_0^*, a_1^*)$ is a local maximum value of f .

The second derivatives of function (15) with respect to a_0 and a_1 are given by

$$f_{a_0 a_0} := \frac{\partial^2 f}{\partial a_0^2} = \frac{2m}{1 + a_1^2}, \quad (22)$$

$$\begin{aligned} f_{a_1 a_1} := \frac{\partial^2 f}{\partial a_1^2} &= \frac{1}{(1 + a_1^2)^3} \left((6a_1^2 - 2) \sum_{i=1}^m (y_i - a_1 x_i - a_0)^2 \right. \\ &\quad \left. - 4(a_1 + a_1^3) \left\{ -2 \sum_{i=1}^m x_i y_i + 2a_1 \sum_{i=1}^m x_i^2 \right. \right. \\ &\quad \left. \left. + 2a_0 \sum_{i=1}^m x_i \right\} + 2(1 + a_1^2)^2 \sum_{i=1}^m x_i^2 \right), \end{aligned} \quad (23)$$

$$f_{a_1 a_0} := \frac{\partial^2 f}{\partial a_1 \partial a_0} = f_{a_0 a_1} := \frac{\partial^2 f}{\partial a_0 \partial a_1} = \frac{2}{(1 + a_1^2)^2} \left(\sum_{i=1}^m x_i + 2a_1 \sum_{i=1}^m y_i - a_1^2 \sum_{i=1}^m x_i - 2ma_1 a_0 \right). \quad (24)$$

Using the second derivative test Definition 1, the parameters a_0 and a_1 that minimize (15) are determined by (20) and (21).

On the other hand, if we consider now a polynomial of degree 2, $f_2(x) = a_0 + a_1x + a_2x^2$. Then, substituting $f_2(x)$ and $f_2'(x)$ into the orthogonality condition (13) we obtain the cubic equation

$$x^3 + \hat{\mu}_1 x^2 + \hat{\mu}_2 x + \hat{\mu}_3 = 0, \quad (25)$$

with

$$\begin{aligned} \hat{\mu}_1 &= \frac{3a_1}{2a_2}, \\ \hat{\mu}_2 &= \frac{1 + a_1^2 + 2a_0a_2 - 2a_2y_i}{2a_2^2}, \\ \hat{\mu}_3 &= \frac{a_1a_0 - x_i - a_1y_i}{2a_2^2}. \end{aligned}$$

The real solution of Equation (25) determines the points $(x_i^*, f_2(x_i^*))$ on the curve $f_2(x)$, where the orthogonal distances for each observation $\|\mathbf{d}_i\|_2$ are determined. Then, solving (25) by Cardano's formulas [38], the real solution for each observation is determined by the formula

$$x_i^* = \sqrt[3]{\frac{-\nu + \sqrt{\nu^2 + \frac{4\omega^3}{27}}}{2}} + \sqrt[3]{\frac{-\nu - \sqrt{\nu^2 + \frac{4\omega^3}{27}}}{2}} - \frac{\hat{\mu}_1}{3}, \quad (26)$$

where

$$\begin{aligned} \nu &= \hat{\mu}_3 + \frac{2}{27}\hat{\mu}_1^3 - \frac{\hat{\mu}_1\hat{\mu}_2}{3}, \\ \omega &= \hat{\mu}_2 - \frac{1}{3}\hat{\mu}_1^2. \end{aligned}$$

Hence, the squared 2-norm of vector $\mathbf{d}_i$ (11) is defined by

$$\|\mathbf{d}_i\|_2^2 = (x_i - x_i^*)^2 + (y_i - f_2(x_i^*))^2.$$

Taking the squared 2-norm of the distances vector $\mathbf{d} = [\|\mathbf{d}_1\|_2 \quad \|\mathbf{d}_2\|_2 \quad \dots \quad \|\mathbf{d}_m\|_2]^\top$, the functional of the sum of the squares of the orthogonal distances for the polynomial of degree 2 is defined by

$$\|\mathbf{d}\|_2^2 = \mathbf{d} \cdot \mathbf{d} = \sum_{i=1}^m \left((x_i - x_i^*)^2 + (y_i - f_2(x_i^*))^2 \right). \quad (27)$$

Notice that the polynomial fitting of degree 2 by OD method is a nonlinear minimization problem.

Considering now a cubic polynomial curve, $f_3(x) = a_0 + a_1x + a_2x^2 + a_3x^3$, then by the orthogonality condition (13), a fifth degree equation is obtained:

$$x^5 + \bar{\mu}_1 x^4 + \bar{\mu}_2 x^3 + \bar{\mu}_3 x^2 + \bar{\mu}_4 x + \bar{\mu}_5 = 0, \quad (28)$$

with

$$\begin{aligned}\bar{\mu}_1 &= \frac{5a_2}{3a_3}, \\ \bar{\mu}_2 &= \frac{4a_1a_3 + 2a_2^2}{3a_3^2}, \\ \bar{\mu}_3 &= \frac{a_1a_2 + a_0a_3 - a_3y_i}{a_3^2}, \\ \bar{\mu}_4 &= \frac{1 + a_1^2 + 2a_0a_2 - 2a_2y_i}{3a_3^2}, \\ \bar{\mu}_5 &= \frac{a_0a_1 - a_1y_i - x_i}{3a_3^2}.\end{aligned}$$

Equation (28) in general cannot be solved analytically; the roots must be obtained with a numerical method [31,39,40]. Once Equation (28) is solved for each observation, the points $(x_i^*, f_3(x_i^*))$ on the curve determine the orthogonal distance vectors $\|\mathbf{d}_i\|_2$, then the squared 2-norm of the distances vector $\mathbf{d} = [\|\mathbf{d}_1\|_2 \quad \|\mathbf{d}_2\|_2 \quad \dots \quad \|\mathbf{d}_m\|_2]^\top$ determines the functional to be minimized:

$$\|\mathbf{d}\|_2^2 = \mathbf{d} \cdot \mathbf{d} = \sum_{i=1}^m \left((x_i - x_i^*)^2 + (y_i - f_3(x_i^*))^2 \right). \quad (29)$$

The general case is when a polynomial curve of degree n is considered, $f_n(x) = a_0 + a_1x + \dots + a_nx^n$. From the orthogonality condition (13), an equation of degree $2n - 1$ is obtained. Since it is of an odd degree, it has at least one real root. Hence the functional of the sum of the squares of orthogonal distances to be minimized is given by

$$\|\mathbf{d}\|_2^2 = \mathbf{d} \cdot \mathbf{d} = \sum_{i=1}^m \left((x_i - x_i^*)^2 + (y_i - f_n(x_i^*))^2 \right). \quad (30)$$

As we can see, there are two numerical problems when the polynomial fitting problem of degree n is solved by the OD method. Firstly, a real-root-finding problem of an equation of degree $2n - 1$ given by the orthogonality condition (13) and then a nonlinear minimization problem of the sum of the squares of the orthogonal distances is given by (30). In this sense, the real-root-finding problem can be solved by the *fzero* routine from MATLAB, whereas the *fminsearch* routine also from MATLAB [41] can be used to solve the nonlinear minimization problem. It is worth mentioning that the polynomial fitting problem of degree 1 can be solved by the analytic formulas (20) and (21), and the second derivative test Definition 1.

Contrasting the OD and TLS Methods

In previous sections, the LS, TLS and OD methods for the polynomial fitting were presented. However, Golub and Van Loan [11] and Van Huffel and Vandewalle [12] mentioned that the TLS method also minimizes the orthogonal distances. In order to review this statement, this section analyzes the parameter estimation by the TLS and OD methods for the most simple polynomial, the straight line $f_1(x) = a_0 + a_1x$.

Applying the OD method, the functional of the sum of the squares of the orthogonal distances for the straight line fitting to be minimized is given by

$$\|\mathbf{d}\|_2^2 = \frac{1}{1 + a_1^2} \sum_{i=1}^m r_i^2, \quad (31)$$

with $r_i = y_i - (a_0 + a_1x_i)$.

On the other hand, the $\mathbf{z}$ TLS solution for the straight line satisfies the eigenvalue–eigenvector equation

$$\Lambda^\top \Lambda \mathbf{z} = \sigma_2^2 \mathbf{z}, \quad (32)$$

where $\Lambda := [\mathbf{A} \ \mathbf{y}]$ and $\mathbf{z} := \begin{bmatrix} \mathbf{p} \\ -1 \end{bmatrix}$. Premultiplying Equation (32) by $\mathbf{z}^\top$ and solving for σ_2^2 , we obtain

$$\sigma_2^2 = \frac{\|\Lambda \mathbf{z}\|_2^2}{\|\mathbf{z}\|_2^2},$$

and substituting Λ and $\mathbf{z}$ results in

$$\sigma_2^2 = \frac{\|\mathbf{A}\mathbf{p} - \mathbf{y}\|_2^2}{\|\mathbf{p}\|_2^2 + 1}. \quad (33)$$

Notice that the numerator of Equation (33) is the same as in Equation (2), then, substituting (2) in (33) reduces to

$$\sigma_2^2 = \frac{1}{\|\mathbf{p}\|_2^2 + 1} \sum_{i=1}^m r_i^2, \quad (34)$$

where $\mathbf{p}$ is the parameters vector. Since the parameters vector for the straight line is $\mathbf{p} = [a_0 \ a_1]^\top$, Equation (34) reduces to

$$\sigma_2^2 = \frac{1}{a_0^2 + a_1^2 + 1} \sum_{i=1}^m r_i^2. \quad (35)$$

Notice that Equations (35) and (31) are not equal, since

$$\frac{1}{1 + a_1^2} \sum_{i=1}^m r_i^2 \neq \frac{1}{a_0^2 + a_1^2 + 1} \sum_{i=1}^m r_i^2,$$

then TLS method does not minimize the orthogonal distances. However, if we consider the straight line without independent coefficient $f_1(x) = a_1 x$, Equations (35) and (31) are equal and both methods minimize the orthogonal distances. Hence the OD and TLS methods are not equivalent in general.

The next section presents the parameter estimation of polynomial functions of degree 1 to 4 with and without an independent coefficient by the LS, TLS and OD methods.

4. Polynomial Fitting of Thermoelectric Voltage by LS, TLS and OD Methods

Considering the observations of temperature and thermoelectric voltage (T_i, E_i) from a R-type thermocouple for the temperature range -50°C to 1064°C with increment of 1°C obtained from [22], this section shows the calibration curve fitting of the R-type thermocouple by polynomials of degree 1 to 4 by LS, TLS and OD methods.

In order to evaluate the predictive performance of the calibration polynomials, the criteria of the extreme values of errors, the average of the absolute errors, the standard deviation and the coefficient of determination R^2 were used [3,4,6].

The error of calibration polynomial is defined as

$$e_i = E_i - \hat{E}_i, \quad (36)$$

where E_i is thermoelectric voltage data and $\hat{E}_i$ is the predicted value of the thermoelectric voltage with the calibration polynomial. The average of the absolute errors is given by

$$e_{ave} = \frac{1}{m} \sum_{i=1}^m |e_i|, \quad (37)$$

where $|e_i|$ is the absolute value of e_i and m is the number of data. The standard deviation of the calibration polynomial is defined by

$$e_{std} = \sqrt{\frac{1}{m-1} \sum_{i=1}^m e_i^2}. \quad (38)$$

The coefficient of determination R^2 is given by

$$R^2 = 1 - \frac{\sum_{i=1}^m (E_i - \hat{E}_i)^2}{\sum_{i=1}^m (E_i - \bar{E})^2}, \quad (39)$$

where $\bar{E}$ is the average of the thermoelectric voltage data.

The calibration polynomial fitting with an independent coefficient is presented below.

4.1. Calibration Polynomial Fitting with Independent Coefficient

The polynomial fitting of functions (40) by LS and TLS methods was solved by Equations (3) and (10), respectively.

$$E(T) = a_0 + \sum_{j=1}^n a_j T^j. \quad (40)$$

On the other hand, the polynomial fitting of degree $n = 1$ by the OD method was carried out applying Algorithm 1, which uses the analytic formulas (20) and (21), and the second derivative test Definition 1. The nonlinear minimization of function (27) for the polynomial fitting of degree $n = 2$ was solved applying the Algorithm 2 which uses the *fminsearch* routine from MATLAB. Finally, the polynomial fitting of functions of degree $n = 3$ and $n = 4$ was solved applying the Algorithm 3 which uses the *fzero* and *fminsearch* routines.

Algorithm 1 Polynomial fitting of degree 1 by OD method

Input Data (x_i, y_i) .

Output Parameters a_0, a_1 .

Step 1. Compute

$$a_1^{(1,2)} = \frac{-\mu_2 \pm \sqrt{\mu_2^2 - 4\mu_1\mu_3}}{2\mu_1},$$

$$a_0^{(1,2)} = \frac{1}{m} \left(\sum_{i=1}^m y_i - a_1^{(1,2)} \sum_{i=1}^m x_i \right).$$

Step 2. Compute $H_f(a_0^*, a_1^*)$ and $f_{a_1 a_1}(a_0^*, a_1^*)$.

Step 3. If $H_f(a_0^*, a_1^*) > 0$ and $f_{a_1 a_1}(a_0^*, a_1^*) > 0$ then (a_0^*, a_1^*) is a local minimum.

Step 4. Output (a_0^*, a_1^*) .

Algorithm 2 Polynomial fitting of degree 2 by OD method

Input Data (x_i, y_i) ; initial approximation of the parameters $P_0 = (a_0, a_1, a_2)$; optimization options: TolFun, TolX, MaxIter, MaxFunEvals.

Output Parameters $P = (a_0, a_1, a_2)$.

function fun(P, x_i, y_i)

$$x_i^* = \sqrt[3]{\frac{-v + \sqrt{v^2 + \frac{4\omega^3}{27}}}{2}} + \sqrt[3]{\frac{-v - \sqrt{v^2 + \frac{4\omega^3}{27}}}{2}} - \frac{\hat{\mu}_1}{3};$$

$$\sum_{i=1}^m \left((x_i - x_i^*)^2 + (y_i - f_2(x_i^*))^2 \right);$$

end function

Step 1. Call *fminsearch*(fun, P_0 , options).

Step 2. Output $P = (a_0, a_1, a_2)$.

Algorithm 3 Polynomial fitting of degree $n > 2$ by OD method

Input Data (x_i, y_i) ; initial approximation of the orthogonality condition solution x_0^* ; initial approximation of the parameters P_0 ; optimization options: TolFun, TolX, MaxIter, MaxFunEvals.

Output Parameters P .

function funZ(P, x_i, y_i, x_0^*)

$$x_i - x + f'_n(x)(y_i - f_n(x));$$

end function

function fun(P, x_i, y_i, x_0^*)

$$x_i^* = \text{fzero}(\text{funZ}, x_0^*, \text{options});$$

$$\sum_{i=1}^m \left((x_i - x_i^*)^2 + (y_i - f_n(x_i^*))^2 \right);$$

end function

Step 1. Call *fminsearch*(fun, P_0 , options).

Step 2. Output P .

Using Algorithms 1–3 and long output display format, the calibration polynomials of degree 1 to 4 by the LS, TLS and OD methods are presented in Tables 1–4, respectively. The performance of the calibration polynomials is presented in Table 5.

Table 1. Calibration polynomial of degree 1.

Parameter	LS	TLS	OD
a_1	0.0106705919592424	0.0108882799524002	0.0106706009250352
a_0	−0.561949764591521	−0.716799304506316	−0.561954310248448

Table 2. Calibration polynomial of degree 2.

Parameter	LS	TLS	OD
a_2	$3.13324263891139 \times 10^{-6}$	$3.12919113815010 \times 10^{-6}$	$3.13325997351179 \times 10^{-6}$
a_1	0.00749348392338622	0.00749870116623730	0.00749346705386636
a_0	−0.0811630813784754	−0.0825736260294239	−0.0811607801463052

Table 3. Calibration polynomial of degree 3.

Parameter	LS	TLS	OD
a_3	$-1.78686899229851 \times 10^{-9}$	$-1.78635025917305 \times 10^{-9}$	$-1.78689809449655 \times 10^{-9}$
a_2	$5.85107037619782 \times 10^{-6}$	$5.85010421131572 \times 10^{-6}$	$5.85111575583941 \times 10^{-6}$
a_1	0.00644876666562105	0.00644929803644297	0.00644874868154519
a_0	−0.0172346997190748	−0.0173148908943473	−0.0172335743482728

Table 4. Calibration polynomial fitting of degree 4.

Parameter	LS	TLS	OD
a_4	$3.06378459910204 \times 10^{-12}$	$3.06378816842890 \times 10^{-12}$	$3.06382882864091 \times 10^{-12}$
a_3	$-8.00022415925578 \times 10^{-9}$	$-8.00023360534995 \times 10^{-9}$	$-8.00031467361401 \times 10^{-9}$
a_2	$9.76012119529912 \times 10^{-6}$	$9.76012992574279 \times 10^{-6}$	$9.76017917221101 \times 10^{-6}$
a_1	0.00567926259968995	0.00567925934996503	0.00567925074682013
a_0	0.000765570482434280	0.000765962049712438	0.000765883225715231

Table 5. Prediction performance of the calibration polynomials.

Polynomial		Criteria				
		$e_{\max}$	$e_{\min}$	e_{ave}	e_{std}	R^2
$n = 1$	LS	0.869479362553643	−0.307193128619244	0.247974185136004	0.295191290947890	0.992673799531263
	TLS	1.03521330212633	−0.256564111788997	0.239056263901570	0.306646708107496	0.992094154347717
	OD	0.869484356500207	−0.307192716192777	0.247973923511759	0.295191290962009	0.992673799530562
$n = 2$	LS	0.222004170950508	−0.0780352716454846	0.0418017281536729	0.0525832990751692	0.999767529752450
	TLS	0.223685706495914	−0.0775891855822355	0.0416607726718407	0.0525863417404143	0.999767502848486
	OD	0.222000982905844	−0.0780392481403069	0.0418020852409217	0.0525832991002195	0.999767529752229
$n = 3$	LS	0.0988219984355953	−0.0380362681738999	0.0198691042767306	0.0239141892875648	0.999951917942603
	TLS	0.0989312384058100	−0.0380051960408253	0.0198639506867380	0.0239142059253218	0.999951917875699
	OD	0.0988198567741218	−0.0380358812609328	0.0198692556077872	0.0239141893000085	0.999951917942553
$n = 4$	LS	0.0317780798401641	−0.0218970177575368	0.00675022233195031	0.00793478730849462	0.999994706507555
	TLS	0.0317775027574607	−0.0218970316206750	0.00675023870006046	0.00793478730961125	0.999994706507553
	OD	0.0317770179203829	−0.0218980118001859	0.00675025296960785	0.00793478731526725	0.999994706507546

Since each set of parameters in Tables 1–4 is distinct, different calibration polynomials were obtained with all methods used. It is important to mention that if we had used a shorter output display format, some parameters estimated would have been the same.

The polynomial fitting without the independent coefficient a_0 is presented below.

4.2. Calibration Polynomial Fitting without Independent Coefficient

The calibration polynomials for the R-type thermocouple without independent coefficient a_0 are defined by

$$E(T) = \sum_{j=1}^n a_j T^j. \quad (41)$$

Using the observations of temperature and thermoelectric voltage for the temperature range -50°C to 1064°C , the calibration polynomial fitting of degrees 1 to 4 without an independent coefficient by the LS, TLS and OD methods are presented in Tables 6–9, respectively. The performance of the calibration polynomials is shown in Table 10.

Table 6. Calibration polynomial of degree 1 without independent coefficient.

Parameter	LS	TLS	OD
a_1	0.00988060792856485	0.00988061279849602	0.00988061279849602

Table 7. Calibration polynomial of degree 2 without independent coefficient.

Parameter	LS	TLS	OD
a_2	$3.36630540516376 \times 10^{-6}$	$3.36630397586626 \times 10^{-6}$	$3.36632169932580 \times 10^{-6}$
a_1	0.00719333493773556	0.00719333615493144	0.00719332224566322

Table 8. Calibration polynomial of degree 3 without independent coefficient.

Parameter	LS	TLS	OD
a_3	$-1.89795367055803 \times 10^{-9}$	$-1.89795171695059 \times 10^{-9}$	$-1.89797754460899 \times 10^{-9}$
a_2	$6.05810804037494 \times 10^{-6}$	$6.05810506932127 \times 10^{-6}$	$6.05814294168891 \times 10^{-6}$
a_1	0.00633478106141306	0.00633478211573470	0.00633476947536766

Table 9. Calibration polynomial of degree 4 without independent coefficient.

Parameter	LS	TLS	OD
a_4	$3.05426851859784 \times 10^{-12}$	$3.05426721773752 \times 10^{-12}$	$3.05430880341339 \times 10^{-12}$
a_3	$-7.97635148760036 \times 10^{-9}$	$-7.97634871709437 \times 10^{-9}$	$-7.97643212780673 \times 10^{-9}$
a_2	$9.73945467428371 \times 10^{-6}$	$9.73945283014663 \times 10^{-6}$	$9.73950412699805 \times 10^{-6}$
a_1	0.00568634618059175	0.00568634655477984	0.00568633723894135

Table 10. Prediction performance of the calibration polynomials without independent coefficient.

Polynomial		Criteria				
		$e_{\max}$	$e_{\min}$	e_{ave}	e_{std}	R^2
$n = 1$	LS	0.848033164006997	−0.565436360926447	0.373819437547787	0.421821855707890	0.985040059251142
	TLS	0.848027982400240	−0.565437933914213	0.373819629178487	0.421821855718038	0.985040059250422
	OD	0.848027982400236	−0.565437933914215	0.373819629178487	0.421821855718038	0.985040059250422
$n = 2$	LS	0.125250983373869	−0.114185743980775	0.0530813308730469	0.0618423117070237	0.999678453972557
	TLS	0.125251047806906	−0.114185887189530	0.0530813028252546	0.0618423117072944	0.999678453972554
	OD	0.125250308034846	−0.114184278342810	0.0530815613672524	0.0618423117402948	0.999678453972211
$n = 3$	LS	0.0753565387608956	−0.0452067871939745	0.0212133821710392	0.0246707283771419	0.999948827614715
	TLS	0.0753565991488130	−0.0452068669060518	0.0212133734955633	0.0246707283772243	0.999948827614714
	OD	0.0753558692210848	−0.0452059328828053	0.0212134738720864	0.0246707283878070	0.999948827614670
$n = 4$	LS	0.0329525392296870	−0.0218315193646372	0.00671767771415101	0.00793902257617295	0.999994700855144
	TLS	0.0329525629038775	−0.0218314997375604	0.00671767650802620	0.00793902257618144	0.999994700855144
	OD	0.0329519581835754	−0.0218324861820349	0.00671769420876913	0.00793902258221761	0.999994700855136

From Table 6, regardless of the output display format type, the estimates for a_1 by the TLS and OD methods were the same. On the other hand, since each set of parameters in Tables 7–9 is distinct, the TLS method also minimizes the orthogonal distances when a polynomial of degree 1 without an independent coefficient is considered.

5. Discussion

The article presents the OD method for the polynomial fitting problem for which the minimization criteria are the sum of the squares of the orthogonal distances from data points to the polynomial curve; hence, the functional obtained is nonlinear, just like the minimization problem. Applying the OD method to fit calibration polynomials of a

R-type thermocouple, different polynomials were obtained compared with the LS and TLS methods. However, if we consider the calibration polynomial of degree 1 without an independent coefficient, the OD and TLS methods obtain the same estimates. An important aspect to consider is the output display format of computations, because if a short output format is used, some results with the OD and TLS methods would be the same. On the other hand, from a practical point of view, the polynomial fitting by the OD method requires computations of roots of equations of degree $2n - 1$ and a nonlinear minimization, so the method can be implemented with different minimization and root finding algorithms.

6. Conclusions

This work presented two classical methods for solving the polynomial fitting problem, the least squares (LS) and total least squares (TLS) methods. Then, using elementary vector geometry, the orthogonal distances (OD) method for solving the polynomial fitting problem of degree n was developed. All methods used in this work have advantages and disadvantages. The LS and TLS solutions are easier to obtain than the OD method, as they are obtained by Equations (3) and (10), respectively. The OD solution for the polynomial fitting of degree $n > 1$ needs to solve two numerical problems. Firstly a real-root finding routine is applied to an equation of degree $2n - 1$ resulting from the orthogonality condition (13); since it is of an odd degree, it has at least one real root. Then a nonlinear minimization routine is applied to the functional of the sum of the squares of the orthogonal distances (30). The polynomial fitting problem of degree 1 by OD method is solved applying the analytic formulas (20), (21) and the second derivative test Definition 1. On the other hand, the LS and TLS methods are ill-conditioned problems [10,12,39] which cause deficient estimates; furthermore, the TLS method has nongeneric problems [35–37]. Since the OD method considers a different minimization criterion than the LS and TLS methods, it represents a different alternative for polynomial fitting. Furthermore, the numerical experiments demonstrated that TLS and OD methods are not equivalent in general, and that the performance criteria of the LS and OD methods are similar. However, from the cases studied, TLS and OD methods obtain the same estimates when the polynomial of degree 1 without an independent coefficient is considered, and therefore, the performance criteria are also equal. It is important to highlight that when the independent coefficient is not considered, the performance criteria are similar to the three methods used. On the other hand, since the minimization of the functional of the OD method is nonlinear and numerical computations of roots are also needed, the implementation of the OD method needs a large computing capacity and efficient algorithms.

Finally, future works include applying the calibration curves obtained in practical applications and solving polynomial fitting problems from different disciplines with the OD method, using different minimization and root-finding algorithms.

Author Contributions: Conceptualization, C.V.-J.; methodology, C.V.-J.; software, L.A.C.-C.; validation, L.A.C.-C. and Y.L.-H.; formal analysis, L.A.C.-C.; investigation, S.I.P.-R. and C.M.M.-V.; data curation, Y.L.-H.; writing-original draft preparation, L.A.C.-C.; writing-review and editing, L.A.C.-C., S.I.P.-R. and C.M.M.-V.; visualization, S.I.P.-R., Y.L.-H. and C.M.M.-V.; supervision, C.V.-J.; project administration, L.A.C.-C.; funding acquisition, C.V.-J., L.A.C.-C., S.I.P.-R., Y.L.-H. and C.M.M.-V. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Secretaría de Investigación y Posgrado SIP IPN under grants 20221223, 20227019, 20221135 and 20220039.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We are greatly indebted to the anonymous referees for their useful suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Bard, Y. *Nonlinear Parameter Estimation*; Academic Press: New York, NY, USA, 1974.
- Lancaster, P.; Salkauskas, K. *Curve and Surface Fitting: An Introduction*; Academic Press: London, UK, 1986.
- Chen, A.; Chen, C. Evaluation of piecewise polynomial equations for two types of thermocouples. *Sensors* **2013**, *13*, 17084–17097. [CrossRef] [PubMed]
- Chen, C.; Weng, Y.K.; Shen, T.C. Performance evaluation of an infrared thermocouple. *Sensors* **2010**, *10*, 10081–10094. [CrossRef] [PubMed]
- Izonin, I.; Tkachenko, R.; Kryvinska, N.; Tkachenko, P. Multiple linear regression based on coefficients identification using non-iterative SGTm neural-like structure. In *International Work-Conference on Artificial Neural Networks*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 467–479.
- Chen, A.; Chen, H.Y.; Chen, C. A software improvement technique for platinum resistance thermometers. *Instruments* **2020**, *4*, 15. [CrossRef]
- Zheng, J.; Hu, G.; Ji, X.; Qin, X. Quintic generalized Hermite interpolation curves: construction and shape optimization using an improved GWO algorithm. *Comput. Appl. Math.* **2022**, *41*, 1–29. [CrossRef]
- Abdulle, A.; Wanner, G. 200 years of least squares method. *Elemente Math.* **2002**, *57*, 45–60. [CrossRef]
- Lawson, C.L.; Hanson, R.J. *Solving Least Squares Problems*; SIAM: Philadelphia, PA, USA, 1995.
- Björck, Å. *Numerical Methods for Least Squares Problems*; SIAM: Philadelphia, PA, USA, 1996.
- Golub, G.H.; Van Loan, C.F. An analysis of the total least squares problem. *SIAM J. Numer. Anal.* **1980**, *17*, 883–893. [CrossRef]
- Van Huffel, S.; Vandewalle, J. *The Total Least Squares Problem: Computational Aspects and Analysis*; SIAM: Philadelphia, PA, USA, 1991.
- Deming, W.E. *Statistical Adjustment of Data*; John Wiley & Sons: New York, NY, USA, 1943.
- Adcock, R.J. A problem in least squares. *Analyst* **1878**, *5*, 53–54. [CrossRef]
- Pearson, K. LIII. On lines and planes of closest fit to systems of points in space. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1901**, *2*, 559–572. [CrossRef]
- Petrás, I.; Podlubny, I. Least squares or least circles? A comparison of classical regression and orthogonal regression. *Chance* **2010**, *23*, 38–42. [CrossRef]
- Scariano, S.M.; Barnett, W., II. Contrasting total least squares with ordinary least squares part I: Basic ideas and results. *Math. Comput. Educ.* **2003**, *37*, 141–158.
- Smith, D.; Pourfarzaneh, M.; Kamel, R. Linear regression analysis by Deming's method. *Clin. Chem.* **1980**, *26*, 1105–1106. [CrossRef] [PubMed]
- Linnet, K. Performance of Deming regression analysis in case of misspecified analytical error ratio in method comparison studies. *Clin. Chem.* **1998**, *44*, 1024–1031. [CrossRef] [PubMed]
- Cantera, L.A.C.; Luna, L.; Vargas-Jarillo, C.; Garrido, R. Parameter estimation of a linear ultrasonic motor using the least squares of orthogonal distances algorithm. In *Proceedings of the 16th International Conference on Electrical Engineering, Computing Science and Automatic Control*, Mexico City, Mexico, 11–13 September 2019.
- Luna, L.; Lopez, K.; Cantera, L.; Garrido, R.; Vargas, C. Parameter estimation and delay-based control of a linear ultrasonic motor. In *Robótica y Computación. Nuevos Avances*; Castro-Liera, I., Cortés-Larinaga, M., Eds.; Instituto Tecnológico de La Paz: La Paz, BCS, México, 2020; pp. 9–15.
- Garrity, K. NIST ITS-90 Thermocouple Database—SRD 60. 2000. Available online: <https://data.nist.gov/od/id/ECBCC1C1302A2ED9E04306570681B10748> (accessed on 28 March 2022).
- Åström, K.J.; Eykhoff, P. System identification—a survey. *Automatica* **1971**, *7*, 123–162. [CrossRef]
- Englezos, P.; Kalogerakis, N. *Applied Parameter Estimation for Chemical Engineers*; CRC Press: Boca Raton, FL, USA, 2000.
- Ding, F.; Shi, Y.; Chen, T. Performance analysis of estimation algorithms of nonstationary ARMA processes. *IEEE Trans. Signal Process.* **2006**, *54*, 1041–1053. [CrossRef]
- Liu, P.; Liu, G. Multi-innovation least squares identification for linear and pseudo-linear regression models. *IEEE Trans. Syst. Man, Cybern. Part Cybern.* **2010**, *40*, 767–778.
- Sun, X.; Ji, J.; Ren, B.; Xie, C.; Yan, D. Adaptive forgetting factor recursive least square algorithm for online identification of equivalent circuit model parameters of a lithium-ion battery. *Energies* **2019**, *12*, 2242. [CrossRef]
- Li, M.; Liu, X. Maximum likelihood hierarchical least squares-based iterative identification for dual-rate stochastic systems. *Int. J. Adapt. Control. Signal Process.* **2021**, *35*, 240–261. [CrossRef]
- Zhang, X.; Ding, F. Adaptive parameter estimation for a general dynamical system with unknown states. *Int. J. Robust Nonlinear Control* **2020**, *30*, 1351–1372. [CrossRef]
- Kang, Z.; Ji, Y.; Liu, X. Hierarchical recursive least squares algorithms for Hammerstein nonlinear autoregressive output-error systems. *Int. J. Adapt. Control. Signal Process.* **2021**, *35*, 2276–2295. [CrossRef]
- Burden, R.; Faires, J.; Burden, A. *Numerical Analysis*; Cengage Learning: Boston, MA, USA, 2016.
- Strang, G. *Introduction to Linear Algebra*; Wellesley-Cambridge Press: Wellesley, MA, USA, 2016.
- Eckart, C.; Young, G. The approximation of one matrix by another of lower rank. *Psychometrika* **1936**, *1*, 211–218. [CrossRef]
- Mirsky, L. Symmetric gauge functions and unitarily invariant norms. *Q. J. Math.* **1960**, *11*, 50–59. [CrossRef]

-
35. Van Huffel, S.; Vandewalle, J. Analysis and solution of the nongeneric total least squares problem. *SIAM J. Matrix Anal. Appl.* **1988**, *9*, 360–372. [[CrossRef](#)]
 36. Van Huffel, S. On the significance of nongeneric total least squares problems. *SIAM J. Matrix Anal. Appl.* **1992**, *13*, 20–35. [[CrossRef](#)]
 37. Van Huffel, S.; Zha, H. The total least squares problem. In *Handbook of Statistics*; Elsevier: Amsterdam, The Netherlands, 1993; Volume 9, pp. 377–408.
 38. Uspensky, J.V. *Theory of Equations*; McGraw-Hill: New York, NY, USA, 1963.
 39. Atkinson, K.E. *An Introduction to Numerical Analysis*; John Wiley & Sons: New York, NY, USA, 2008.
 40. Dennis, J.E., Jr.; Schnabel, R.B. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*; SIAM: Philadelphia, PA, USA, 1996.
 41. MathWorks. *Optimization Toolbox User's Guide*; MathWorks: Natick, MA, USA, 2020.