Using Parameter Elimination to Solve Discrete Linear Chebyshev Approximation Problems

Nikolai Krivulin [ORCID]

Faculty of Mthematics and Mechanics, St. Petersburg State University, Universitetskaya Emb. 7/9, 199034 St. Petersburg, Russia; nkk@math.spbu.ru

Abstract: We consider discrete linear Chebyshev approximation problems in which the unknown parameters of linear function are fitted by minimizing the least maximum absolute deviation of errors. Such problems find application in the solution of overdetermined systems of linear equations that appear in many practical contexts. The least maximum absolute deviation estimator is used in regression analysis in statistics when the distribution of errors has bounded support. To derive a direct solution of the problem, we propose an algebraic approach based on a parameter elimination technique. As a key component of the approach, an elimination lemma is proved to handle the problem by reducing it to a problem with one parameter eliminated, together with a box constraint imposed on this parameter. We demonstrate the application of the lemma to the direct solution of linear regression problems with one and two parameters. We develop a procedure to solve multidimensional approximation (multiple linear regression) problems in a finite number of steps. The procedure follows a method that comprises two phases: backward elimination and forward substitution of parameters. We describe the main components of the procedure and estimate its computational complexity. We implement symbolic computations in MATLAB to obtain exact solutions for two numerical examples.

Keywords: discrete linear Chebyshev approximation; minimax problem; variable elimination; direct solution; multiple linear regression; least maximum absolute deviation estimator

MSC: 41A50; 90C47; 62J05

1. Introduction

Discrete linear Chebyshev (minimax) approximation problems where the errors of fitting the unknown parameters are measured by the Chebyshev (max, infinity, uniform or L_∞) norm are of theoretical interest and practical importance in many areas of science and engineering. Application of the Chebyshev norm leads to the least maximum absolute deviation of errors as the approximation criterion, and dates back to Laplace's classical work [1] (book 3, chap. V, §39) (see also [2,3]).

An important area of applications of the discrete linear Chebyshev approximation is the solution of overdetermined systems of linear equations [4–6] that appear in many practical contexts. The least maximum absolute deviation estimator is widely used in regression analysis in statistics when the distribution of errors has bounded support. Specifically, the Chebyshev estimator is known to be a maximum likelihood estimator if the error distribution is uniform [5,7–9]. Moreover, this estimator can be useful even if errors are not uniform, but controlled in some way, and small, relative to the observed values. Examples of applications include problems in nuclear physics [10,11], parameter estimation of dynamic systems [12,13], statistical machine learning [14,15], and finance [16].

To solve the Chebyshev approximation problem, a number of approaches are known which apply various iterative computational procedures to find numerical solutions (see a comprehensive

overview of the algorithmic solutions given by [8,17–19]). For instance, the approximation problems under consideration can be reduced to linear programs and then solved numerically by computational algorithms available in the linear programming, such as the simplex algorithm and its variations. For linear programming solutions and other related algorithms, one can consult early works [4,5,7,20–24], as well as more recent publications [25–28].

Along with existing iterative algorithms that find use in applications, direct analytical solutions of the linear Chebyshev approximation problem are also of interest as an essential instrument of formal analysis and treatment of the problem. A useful algebraic approach to derive direct solutions of problems that involve minimizing the Chebyshev distance is proposed in [29–31]. The approach offers complete solutions of the problems in the framework of tropical (idempotent) algebra, which deals with algebraic systems with idempotent operations. The solutions obtained in terms of tropical algebra are then represented in the usual form, ready for computation.

In this paper, we reshape and adjust algebraic techniques implemented in the above-mentioned approach to develop a direct solution of the discrete linear Chebyshev approximation problem in terms of conventional algebra. As a key component of the proposed method, an elimination lemma is proved that allows us to handle the problem by reducing it to a problem with one unknown parameter eliminated and a box constraint imposed on this parameter. To provide illuminating but not cumbersome examples of the application of the lemma, we derive direct solutions of problems of low dimension, formulated as linear regression problems with one and two parameters.

Furthermore, we construct a procedure to solve multidimensional approximation (multiple linear regression) problems. The procedure is based on a direct solution method that comprises two phases: backward elimination and forward determination (substitution) of the unknown parameters. The direct solution can supplement and complement the existing iterative procedures and becomes of particular interest when, for one reason or another, the use of iterative algorithms appears to be inappropriate or inadequate. We estimate computational complexity and memory requirements of the procedure, and implement symbolic computations in the MATLAB environment to obtain exact solutions for two illustrative numerical examples.

The rest of the paper is organized as follows. In Section 2, we formulate the approximation problem of interest in both scalar and vector form. Section 3 presents the main result, which offers a reduction step to separate the problem into a problem of lower dimension by eliminating an unknown parameter, and a box constraint for this parameter. We apply the obtained result in Section 4 to derive direct explicit solutions for linear regression problems with one and two unknown parameters. In Section 5, we describe a computational procedure to solve linear approximation problems of arbitrary dimension and discuss its computational complexity. In Section 6, software implementation is discussed and numerical examples are given. Section 7 presents concluding remarks.

2. Linear Chebyshev Approximation Problem

We start with an appropriate notation, preliminary assumptions, and formal representation of the discrete linear Chebyshev approximation problem under study.

Suppose that, given $X_{ij}, Y_i \in \mathbb{R}$ for all $i = 1, \ldots, M$ and $j = 1, \ldots, N$, where M and N are positive integers, we need to find the unknown parameters $\theta_j \in \mathbb{R}$ for all $j = 1, \ldots, N$ that solve the minimax problem

$$\min_{\theta_1, \ldots, \theta_N} \quad \max_{1 \leq i \leq M} \left| \sum_{j=1}^{N} X_{ij} \theta_j - Y_i \right|. \tag{1}$$

Without loss of generality, we assume that for each $j = 1, \ldots, N$, there exists at least one i, such that $X_{ij} \neq 0$. Otherwise, if $X_{ij} = 0$ for some j and all i, the parameter θ_j does not affect the objective function, and thus can be removed.

Note that we can represent problem (1) in vector form by introducing the matrix and column vectors

$$\boldsymbol{X} = (X_{ij}), \qquad \boldsymbol{Y} = (Y_i), \qquad \boldsymbol{\theta} = (\theta_j).$$

With this matrix-vector notation and the Chebyshev norm defined for any vector $\boldsymbol{V} = (V_i)$ as

$$\|\boldsymbol{V}\| = \max_i |V_i|,$$

the approximation problem takes the form

$$\min_{\boldsymbol{\theta}} \quad \|\boldsymbol{X}\boldsymbol{\theta} - \boldsymbol{Y}\|.$$

To solve problem (1), we first show that the problem can be reduced to a problem of the same form, but with one unknown parameter fewer.

3. Elimination Lemma

The next result offers a reduction approach to the problem, which provides the basis for the proposed solution.

Lemma 1. *Solving problem* (1) *is equivalent to solving the problem*

$$\min_{\theta_1,\dots,\theta_{N-1}} \quad \max_{\substack{1 \le i < k \le M \\ |X_{iN}| + |X_{kN}| \ne 0}} \left| \sum_{j=1}^{N-1} \frac{X_{ij} X_{kN} - X_{kj} X_{iN}}{|X_{iN}| + |X_{kN}|} \theta_j - \frac{Y_i X_{kN} - Y_k X_{iN}}{|X_{iN}| + |X_{kN}|} \right| \tag{2}$$

together with the inequality

$$\max_{\substack{1 \le i \le M \\ X_{iN} \ne 0}} \left(-\frac{\mu}{|X_{iN}|} - \sum_{j=1}^{N-1} \frac{X_{ij}}{X_{iN}} \theta_j + \frac{Y_i}{X_{iN}} \right) \le \theta_N \le \min_{\substack{1 \le i \le M \\ X_{iN} \ne 0}} \left(\frac{\mu}{|X_{iN}|} - \sum_{j=1}^{N-1} \frac{X_{ij}}{X_{iN}} \theta_j + \frac{Y_i}{X_{iN}} \right), \tag{3}$$

where μ is the minimum of the objective function in problem (2), *and the empty sums are defined to be zero.*

Proof. To examine problem (1), we first introduce an auxiliary unknown parameter λ to rewrite the problem as follows:

$$\min_{\theta_1,\dots,\theta_N} \quad \lambda,$$

$$\text{s.t.} \quad \max_{1 \le i \le M} \left| \sum_{j=1}^N X_{ij}\theta_j - Y_i \right| \le \lambda.$$

Note that the inequality constraint is readily represented as the system of inequalities

$$\lambda \ge -\sum_{j=1}^N X_{ij}\theta_j + Y_i,$$

$$\lambda \ge \sum_{j=1}^N X_{ij}\theta_j - Y_i, \qquad i = 1,\dots,M;$$

which, in particular, puts the problem into the form of a linear program.

Next, we continue rearrangement by solving for θ_N those inequalities in which $X_{iN} \ne 0$ to write

$$-\theta_N \geq \frac{\lambda}{X_{iN}} + \sum_{j=1}^{N-1} \frac{X_{ij}}{X_{iN}}\theta_j - \frac{Y_i}{X_{iN}},$$

$$\theta_N \geq \frac{\lambda}{X_{iN}} - \sum_{j=1}^{N-1} \frac{X_{ij}}{X_{iN}}\theta_j + \frac{Y_i}{X_{iN}}, \qquad X_{iN} < 0;$$

$$\theta_N \geq -\frac{\lambda}{X_{iN}} - \sum_{j=1}^{N-1} \frac{X_{ij}}{X_{iN}}\theta_j + \frac{Y_i}{X_{iN}},$$

$$-\theta_N \geq -\frac{\lambda}{X_{iN}} + \sum_{j=1}^{N-1} \frac{X_{ij}}{X_{iN}}\theta_j - \frac{Y_i}{X_{iN}}, \qquad X_{iN} > 0; \qquad i = 1,\ldots,M.$$

Coupling the inequalities with common left-hand sides and adding the inequalities for $X_{iN} = 0$ yield

$$-\theta_N \geq -\frac{\lambda}{|X_{iN}|} + \sum_{j=1}^{N-1} \frac{X_{ij}}{X_{iN}}\theta_j - \frac{Y_i}{X_{iN}},$$

$$\theta_N \geq -\frac{\lambda}{|X_{iN}|} - \sum_{j=1}^{N-1} \frac{X_{ij}}{X_{iN}}\theta_j + \frac{Y_i}{X_{iN}}, \qquad X_{iN} \neq 0;$$

$$\lambda \geq \left| \sum_{j=1}^{N-1} X_{ij}\theta_j - Y_i \right|, \qquad X_{iN} = 0; \qquad i = 1,\ldots,M.$$

By combining these inequalities for all $i = 1,\ldots,M$, we obtain

$$-\theta_N \geq \max_{\substack{1 \leq i \leq M \\ X_{iN} \neq 0}} \left(-\frac{\lambda}{|X_{iN}|} + \sum_{j=1}^{N-1} \frac{X_{ij}}{X_{iN}}\theta_j - \frac{Y_i}{X_{iN}} \right),$$

$$\theta_N \geq \max_{\substack{1 \leq i \leq M \\ X_{iN} \neq 0}} \left(-\frac{\lambda}{|X_{iN}|} - \sum_{j=1}^{N-1} \frac{X_{ij}}{X_{iN}}\theta_j + \frac{Y_i}{X_{iN}} \right), \qquad (4)$$

$$\lambda \geq \max_{\substack{1 \leq i \leq M \\ X_{iN} = 0}} \left| \sum_{j=1}^{N-1} X_{ij}\theta_j - Y_i \right|.$$

The first two inequalities at (4) result in the double inequality

$$\max_{\substack{1 \leq i \leq M \\ X_{iN} \neq 0}} \left(-\frac{\lambda}{|X_{iN}|} - \sum_{j=1}^{N-1} \frac{X_{ij}}{X_{iN}}\theta_j + \frac{Y_i}{X_{iN}} \right) \leq \theta_N \leq -\max_{\substack{1 \leq i \leq M \\ X_{iN} \neq 0}} \left(-\frac{\lambda}{|X_{iN}|} + \sum_{j=1}^{N-1} \frac{X_{ij}}{X_{iN}}\theta_j - \frac{Y_i}{X_{iN}} \right).$$

After replacing λ by μ that denotes the minimum of the objective function and using the equality $\max(a,b) = -\min(-a,-b)$ to change from max to min in the right-hand side, the double inequality takes the form of (3).

The above double inequality defines a nonempty set of values for the unknown θ_N if, and only if, the following condition holds:

$$\max_{\substack{1 \leq i \leq M \\ X_{iN} \neq 0}} \left(-\frac{\lambda}{|X_{iN}|} - \sum_{j=1}^{N-1} \frac{X_{ij}}{X_{iN}}\theta_j + \frac{Y_i}{X_{iN}} \right) + \max_{\substack{1 \leq i \leq M \\ X_{iN} \neq 0}} \left(-\frac{\lambda}{|X_{iN}|} + \sum_{j=1}^{N-1} \frac{X_{ij}}{X_{iN}}\theta_j - \frac{Y_i}{X_{iN}} \right) \leq 0,$$

which is readily rearranged in the form of the inequality

$$\max_{\substack{1 \le i,k \le M \\ X_{iN}, X_{kN} \ne 0}} \left(-\frac{|X_{iN}| + |X_{kN}|}{|X_{iN}||X_{kN}|}\lambda - \sum_{j=1}^{N-1} \frac{X_{ij}X_{kN} - X_{kj}X_{iN}}{X_{iN}X_{kN}}\theta_j + \frac{Y_iX_{kN} - Y_kX_{iN}}{X_{iN}X_{kN}} \right) \le 0.$$

This inequality is equivalent to the system of inequalities

$$-\frac{|X_{iN}| + |X_{kN}|}{|X_{iN}||X_{kN}|}\lambda - \sum_{j=1}^{N-1} \frac{X_{ij}X_{kN} - X_{kj}X_{iN}}{X_{iN}X_{kN}}\theta_j + \frac{Y_iX_{kN} - Y_kX_{iN}}{X_{iN}X_{kN}} \le 0, \qquad X_{iN}, X_{kN} \ne 0; \qquad 1 \le i,k \le M.$$

By solving these inequalities for λ, we obtain the system

$$\lambda \ge -\frac{|X_{iN}||X_{kN}|}{X_{iN}X_{kN}} \left(\sum_{j=1}^{N-1} \frac{X_{ij}X_{kN} - X_{kj}X_{iN}}{|X_{iN}| + |X_{kN}|}\theta_j - \frac{Y_iX_{kN} - Y_kX_{iN}}{|X_{iN}| + |X_{kN}|} \right), \qquad X_{iN}, X_{kN} \ne 0; \qquad 1 \le i,k \le M.$$

We now note that interchanging the indices i and k in the differences

$$X_{ij}X_{kN} - X_{kj}X_{iN}, \qquad Y_iX_{kN} - Y_kX_{iN}$$

changes the sign of these differences, and hence, the sign of the entire right-hand side of each inequality in the system. As a result, for every pair of indices i and k, the system includes both the inequality

$$\lambda \ge -\frac{|X_{iN}||X_{kN}|}{X_{iN}X_{kN}} \left(\sum_{j=1}^{N-1} \frac{X_{ij}X_{kN} - X_{kj}X_{iN}}{|X_{iN}| + |X_{kN}|}\theta_j - \frac{Y_iX_{kN} - Y_kX_{iN}}{|X_{iN}| + |X_{kN}|} \right),$$

and the inequality

$$\lambda \ge -\frac{|X_{kN}||X_{iN}|}{X_{kN}X_{iN}} \left(\sum_{j=1}^{N-1} \frac{X_{kj}X_{iN} - X_{ij}X_{kN}}{|X_{kN}| + |X_{iN}|}\theta_j - \frac{Y_kX_{iN} - Y_iX_{kN}}{|X_{kN}| + |X_{iN}|} \right)$$

$$= \frac{|X_{iN}||X_{kN}|}{X_{iN}X_{kN}} \left(\sum_{j=1}^{N-1} \frac{X_{ij}X_{kN} - X_{kj}X_{iN}}{|X_{iN}| + |X_{kN}|}\theta_j - \frac{Y_iX_{kN} - Y_kX_{iN}}{|X_{iN}| + |X_{kN}|} \right).$$

After coupling the paired inequalities and considering that the equality $|X_{iN}X_{kN}| = |X_{iN}||X_{kN}|$ is valid, we rearrange the system as follows:

$$\lambda \ge \left| \sum_{j=1}^{N-1} \frac{X_{ij}X_{kN} - X_{kj}X_{iN}}{|X_{iN}| + |X_{kN}|}\theta_j - \frac{Y_iX_{kN} - Y_kX_{iN}}{|X_{iN}| + |X_{kN}|} \right|, \qquad X_{iN}, X_{kN} \ne 0; \qquad 1 \le i,k \le M.$$

Furthermore, we combine the inequalities for all $1 \le i,k \le M$ and add the last inequality at (4) to replace the condition $X_{iN}, X_{kN} \ne 0$ by that in the form $|X_{iN}| + |X_{kN}| \ne 0$ and rewrite the system as one inequality

$$\lambda \ge \max_{\substack{1 \le i,k \le M \\ |X_{iN}| + |X_{kN}| \ne 0}} \left| \sum_{j=1}^{N-1} \frac{X_{ij}X_{kN} - X_{kj}X_{iN}}{|X_{iN}| + |X_{kN}|}\theta_j - \frac{Y_iX_{kN} - Y_kX_{iN}}{|X_{iN}| + |X_{kN}|} \right|.$$

We now observe that the term under the max operator is invariant under permutation of the indices i and k, and is equal to zero if $i = k$. Therefore, we can reduce the set of indices defined as $1 \le i,k \le M$ by that given by the condition $1 \le i < k \le M$ and represent the lower bound on λ as

$$\lambda \ge \max_{\substack{1 \le i < k \le M \\ |X_{iN}| + |X_{kN}| \ne 0}} \left| \sum_{j=1}^{N-1} \frac{X_{ij}X_{kN} - X_{kj}X_{iN}}{|X_{iN}| + |X_{kN}|}\theta_j - \frac{Y_iX_{kN} - Y_kX_{iN}}{|X_{iN}| + |X_{kN}|} \right|.$$

Since the minimum of λ is bounded from below by the expression on the right-hand side, we need to find the minimum of this expression with respect to $\theta_1, \ldots, \theta_{N-1}$, which leads to solving problem (2). □

To conclude this section, we note that the reduced problem at (2) has the same general form as problem (1) with the parameter θ_N eliminated. This offers a potential for solving the problem under study by recurrent implementation of Lemma 1. We discuss application of the lemma to derive direct solutions of problems of low dimension and to develop a recursive procedure to solve problems of arbitrary dimension in what follows.

4. Solution of One- and Two-Parameter Regression Problems

We now apply the obtained result to derive direct, exact solutions to regression problems with one and two parameters. These solutions can be directly extended to problems with more parameters, which leads, however, to more complicated and bulky expressions, not presented here to save space.

We start with one-parameter simple linear regression problems, which have well-known solutions, and then find a complete solution for a two-parameter linear regression problem.

4.1. One-Parameter Linear Regression Problems

Let us suppose that, for given explanatory (independent) variables $X_i \in \mathbb{R}$ and response (dependent) variables $Y_i \in \mathbb{R}$, where $i = 1, \ldots, M$, we find the unknown regression parameter $\theta \in \mathbb{R}$ that achieves the minimum

$$\min_{\theta} \quad \max_{1 \leq i \leq M} |X_i \theta - Y_i|. \tag{5}$$

To solve the problem, we directly apply Lemma 1 with $N = 1$. Elimination of the empty sums in (2) and (3), and substitution $X_{i1} = X_i$ for all $i = 1, \ldots, M$ and $\theta_1 = \theta$ yield the next results.

Proposition 1. *The minimum error in problem* (5) *is equal to*

$$\mu = \max_{\substack{1 \leq i < k \leq M \\ |X_i| + |X_k| \neq 0}} \frac{|Y_i X_k - Y_k X_i|}{|X_i| + |X_k|},$$

and all solutions of the problem are given by the condition

$$\max_{\substack{1 \leq i \leq M \\ X_i \neq 0}} \left(-\frac{\mu}{|X_i|} + \frac{Y_i}{X_i} \right) \leq \theta \leq \min_{\substack{1 \leq i \leq M \\ X_i \neq 0}} \left(\frac{\mu}{|X_i|} + \frac{Y_i}{X_i} \right).$$

We now consider a special case of problem (5) in the form

$$\min_{\theta} \quad \max_{1 \leq i \leq M} |\theta - Y_i|. \tag{6}$$

To handle the problem, we set $X_i = 1$ for all $i = 1, \ldots, M$ in the expressions obtained in Proposition 1. Since $|X_i| + |X_k| = 2 \neq 0$, the minimum error in problem (6) becomes

$$\mu = \max_{1 \leq i < k \leq M} |Y_i - Y_k|/2 = \max_{1 \leq i,k \leq M} |Y_i - Y_k|/2 = \max_{1 \leq i \leq M} Y_i/2 - \min_{1 \leq i \leq M} Y_i/2.$$

The solution θ is given by the condition

$$-\mu + \max_{1 \leq i \leq M} Y_i \leq \theta \leq \mu + \min_{1 \leq i \leq M} Y_i,$$

which, after substitution of the above expression for μ, leads to the unique result

$$\theta = \max_{1 \le i \le M} Y_i / 2 + \min_{1 \le i \le M} Y_i / 2.$$

4.2. Two-Parameter Linear Regression Problem

We now turn to two-parameter problems, which can be solved by twofold application of Lemma 1. To avoid cumbersome calculations, we concentrate on a special case in which, given variables $X_i, Y_i \in \mathbb{R}$ for all $i = 1, \ldots, M$, our aim is to find the parameters $\theta_1, \theta_2 \in \mathbb{R}$ to achieve

$$\min_{\theta_1, \theta_2} \ \max_{1 \le i \le M} |\theta_1 + X_i \theta_2 - Y_i|. \tag{7}$$

Proposition 2. *The minimum error in problem* (7) *is equal to*

$$\mu = \max_{\substack{1 \le i < k \le M \\ |X_i| + |X_k| \ne 0}} \max_{\substack{1 \le p < r \le M \\ X_r \ne X_p}} \frac{|(Y_i X_k - Y_k X_i)(X_r - X_p) - (Y_p X_r - Y_r X_p)(X_k - X_i)|}{(|X_i| + |X_k|)|X_r - X_p| + (|X_p| + |X_r|)|X_k - X_i|}, \tag{8}$$

and all solutions of the problem are given by the conditions

$$\max_{\substack{1 \le i < k \le M \\ X_k \ne X_i}} \left(-\frac{|X_i| + |X_k|}{|X_k - X_i|} \mu + \frac{Y_i X_k - Y_k X_i}{X_k - X_i} \right) \le \theta_1 \le \min_{\substack{1 \le i < k \le M \\ X_k \ne X_i}} \left(\frac{|X_i| + |X_k|}{|X_k - X_i|} \mu + \frac{Y_i X_k - Y_k X_i}{X_k - X_i} \right), \tag{9}$$

$$\max_{\substack{1 \le i \le M \\ X_i \ne 0}} \left(-\frac{\mu}{|X_i|} - \frac{\theta_1}{X_i} + \frac{Y_i}{X_i} \right) \le \theta_2 \le \min_{\substack{1 \le i \le M \\ X_i \ne 0}} \left(\frac{\mu}{|X_i|} - \frac{\theta_1}{X_i} + \frac{Y_i}{X_i} \right). \tag{10}$$

Proof. We apply Lemma 1 with $N = 2$, where we take $X_{i1} = 1$ and $X_{i2} = X_i$ for all $i = 1, \ldots, M$. As a result, problem (7) reduces to the one-parameter problem

$$\min_{\theta_1} \ \max_{\substack{1 \le i < k \le M \\ |X_i| + |X_k| \ne 0}} \left| \frac{X_k - X_i}{|X_i| + |X_k|} \theta_1 - \frac{Y_i X_k - Y_k X_i}{|X_i| + |X_k|} \right|$$

and box constraint for θ_2 in the form of the double inequality at (10), where μ is the minimum in the one-parameter problem.

We note that the objective function in the problem does not change if we replace the condition $1 \le i < k \le M$ for indices over which the maximum is taken, by the extended condition $1 \le i, k \le M$.

In a similar way as in Lemma 1, we first represent the one-parameter problem under consideration as

$$\min_{\theta_1} \ \lambda,$$

$$\text{s.t.} \quad \max_{\substack{1 \le i, k \le M \\ |X_i| + |X_k| \ne 0}} \left| \frac{X_k - X_i}{|X_i| + |X_k|} \theta_1 - \frac{Y_i X_k - Y_k X_i}{|X_i| + |X_k|} \right| \le \lambda.$$

The inequality constraint in this problem is equivalent to the system of inequalities

$$\lambda \ge -\frac{X_k - X_i}{|X_i| + |X_k|} \theta_1 + \frac{Y_i X_k - Y_k X_i}{|X_i| + |X_k|},$$

$$\lambda \ge \frac{X_k - X_i}{|X_i| + |X_k|} \theta_1 - \frac{Y_i X_k - Y_k X_i}{|X_i| + |X_k|}, \qquad |X_i| + |X_k| \ne 0; \qquad 1 \le i, k \le M.$$

After solving the inequalities for θ_1, we rewrite the system as

$$-\theta_1 \geq -\frac{|X_i| + |X_k|}{|X_k - X_i|}\lambda - \frac{Y_i X_k - Y_k X_i}{X_k - X_i},$$

$$\theta_1 \geq -\frac{|X_i| + |X_k|}{|X_k - X_i|}\lambda + \frac{Y_i X_k - Y_k X_i}{X_{i1} X_k - X_{k1} X_i}, \qquad X_k \neq X_i;$$

$$\lambda \geq \frac{|Y_i X_k - Y_k X_i|}{|X_i| + |X_k|}, \qquad X_k = X_i; \qquad |X_i| + |X_k| \neq 0; \qquad 1 \leq i,k \leq M.$$

By combining these inequalities, we obtain

$$-\theta_1 \geq \max_{\substack{1 \leq i,k \leq M \\ |X_i| + |X_k| \neq 0 \\ X_k \neq X_i}} \left(-\frac{|X_i| + |X_k|}{|X_k - X_i|}\lambda - \frac{Y_i X_k - Y_k X_i}{X_k - X_i} \right),$$

$$\theta_1 \geq \max_{\substack{1 \leq i,k \leq M \\ |X_i| + |X_k| \neq 0 \\ X_k \neq X_i}} \left(-\frac{|X_i| + |X_k|}{|X_k - X_i|}\lambda + \frac{Y_i X_k - Y_k X_i}{X_k - X_i} \right), \qquad (11)$$

$$\lambda \geq \max_{\substack{1 \leq i,k \leq M \\ |X_i| + |X_k| \neq 0 \\ X_k = X_i}} \frac{|Y_i X_k - Y_k X_i|}{|X_i| + |X_k|} = \max_{\substack{1 \leq i,k \leq M \\ |X_i| \neq 0}} \frac{|Y_i - Y_k|}{2}.$$

The first two inequalities yield the double inequality

$$\max_{\substack{1 \leq i,k \leq M \\ |X_i| + |X_k| \neq 0 \\ X_k \neq X_i}} \left(-\frac{|X_i| + |X_k|}{|X_k - X_i|}\lambda + \frac{Y_i X_k - Y_k X_i}{X_k - X_i} \right) \leq \theta_1 \leq \min_{\substack{1 \leq i,k \leq M \\ |X_i| + |X_k| \neq 0 \\ X_k \neq X_i}} \left(\frac{|X_i| + |X_k|}{|X_k - X_i|}\lambda + \frac{Y_i X_k - Y_k X_i}{X_k - X_i} \right).$$

Since, under the condition $X_k \neq X_i$, the condition $|X_i| + |X_k| \neq 0$ holds as well, we exclude the latter one. Observing that the terms under the max and min operators are invariant under permutation of i and k, we adjust the condition on indices to write the box constraint for θ_1 as (9).

The feasibility condition for the box constraint to be valid for θ_1 takes the form of the inequality

$$\max_{\substack{1 \leq i,k \leq M \\ X_k \neq X_i}} \max_{\substack{1 \leq p,r \leq M \\ X_r \neq X_p}} \left(-\left(\frac{|X_i| + |X_k|}{|X_k - X_i|} + \frac{|X_p| + |X_r|}{|X_r - X_p|} \right)\lambda + \frac{Y_i X_k - Y_k X_i}{X_k - X_i} - \frac{Y_p X_r - Y_r X_p}{X_r - X_p} \right) \leq 0.$$

As before, we represent this inequality as the system of inequalities for each i,k and p,r, and then solve these inequalities for λ. After combining the solutions back into one inequality and adding the last inequality at (11), we obtain

$$\lambda \geq \max_{\substack{1 \leq i,k \leq M \\ |X_i| + |X_k| \neq 0}} \max_{\substack{1 \leq p,r \leq M \\ X_r \neq X_p}} \frac{|(Y_i X_k - Y_k X_i)(X_r - X_p) - (Y_p X_r - Y_r X_p)(X_k - X_i)|}{(|X_i| + |X_k|)|X_r - X_p| + (|X_p| + |X_r|)|X_k - X_i|},$$

where the expression on the right-hand side determines the minimum μ.

Finally, we note that the fractional term under maximization is invariant with respect to interchanging the indices i and k, as well as p and r. Therefore, we can replace the conditions $1 \leq i,k \leq M$ and $1 \leq p,r \leq M$ by the conditions $1 \leq i < k \leq M$ and $1 \leq p < r \leq M$, which yields the representation for μ in the form of (8).　□

5. General Solution Procedure

We now use Lemma 1 to derive a complete solution of problem (1) by performing a series of reduction steps, each eliminating an unknown parameter in the problem and determining a box constraint for this parameter. We observe that the elimination of a parameter from the objective

function as described in Lemma 1 leaves the general form of the problem unchanged. Therefore, we can repeat the elimination over and over again until the function has no more parameters, and thus becomes a constant that shows the minimum of the objective function in the initial problem.

At the same time, together with the elimination of a parameter from the objective function, Lemma 1 offers a box constraint for this parameter, represented in terms of those parameters which are retained in the function. We see that the last constraint does not depend on any other parameters, and thus is directly given by a double inequality with constant values on both sides. As a result, we can take the box constraints in the order from the last constraint to the first, which yields a system of double inequalities that completely determines the solution set of the problem.

We are in a position to describe the solution procedure formally in more detail. The procedure follows a direct solution method that examines the unknown parameters in reversal order, starting from the parameter θ_N and going backward to the parameter θ_1. Let n be the number of parameters in the objective function in the current step of the procedure.

Initially, we take $n = N$ and set $M_n = M$. For all $i = 1, \ldots, M_n$ and $j = 1, \ldots, n$, we define

$$X_{ij}^n = X_{ij}, \qquad Y_i^n = Y_i.$$

We also introduce the matrix-vector notation

$$
X_n = \begin{pmatrix} X_{11}^n & \cdots & X_{1n}^n \\ \vdots & \ddots & \vdots \\ X_{M_n,1}^n & \cdots & X_{M_n,n}^n \end{pmatrix}, \qquad
Y_n = \begin{pmatrix} Y_1^n \\ \vdots \\ Y_{M_n}^n \end{pmatrix}, \qquad
\theta_n = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_n \end{pmatrix}.
$$

For each $n = N, N-1, \ldots, 1$, the procedure produces a two-fold outcome: the reduction of the current problem by eliminating an unknown parameter, and the derivation of a box constraint for the eliminated parameter.

5.1. Elimination of Parameters

Assuming that the norm sign in what follows stands for the Chebyshev norm, we start with eliminating the parameter θ_n from the problem

$$\min_{\theta_1, \ldots, \theta_n} \quad \max_{1 \le i \le M_n} \left| \sum_{j=1}^{n} X_{ij}^n \theta_j - Y_i^n \right| = \min_{\theta_n} \quad \| X_n \theta_n - Y_n \|.$$

It follows from Lemma 1 that, as a result of this elimination, the problem reduces, if $n > 1$, to the problem

$$\min_{\theta_1, \ldots, \theta_{n-1}} \quad \max_{1 \le i \le M_{n-1}} \left| \sum_{j=1}^{n-1} X_{ij}^{n-1} \theta_j - Y_i^{n-1} \right| = \min_{\theta_{n-1}} \quad \| X_{n-1} \theta_{n-1} - Y_{n-1} \|,$$

or degenerates, if $n = 1$, to the constant

$$\max_{1 \le i \le M_0} \left| Y_i^0 \right| = \| Y_0 \|.$$

We now exploit the representation of problem (2) to establish formulas of recalculating the objective function when changing to the reduced problem.

First, we consider the condition for indices in (2), which takes the form $1 \le i < k \le M_n$. We assume that the pairs of indices (i, k) defined by the condition are listed in the order of the sequence

$$(1, 2), \ldots, (1, M_n), (2, 3), \ldots, (2, M_n), \ldots, (M_n - 1, M_n).$$

It is not difficult to verify by direct substitution that each fixed pair (i, k) in this sequence has the number (index) calculated as

$$M_n(i-1) - i(i-1)/2 + k - 1.$$

Furthermore, we use (2) to define, for all i and k such that $1 \leq i < k \leq M_n$ and for all $j = 1, \ldots, n-1$, the recurrent formulas

$$
\begin{aligned}
X^{n-1}_{M_n(i-1)-i(i-1)/2+k-1,j} &= \begin{cases} \dfrac{X^n_{ij} X^n_{kn} - X^n_{kj} X^n_{in}}{|X^n_{in}| + |X^n_{kn}|}, & \text{if } |X^n_{in}| + |X^n_{kn}| \neq 0; \\ 0 & \text{otherwise;} \end{cases} \\[2ex]
Y^{n-1}_{M_n(i-1)-i(i-1)/2+k-1} &= \begin{cases} \dfrac{Y^n_i X^n_{kn} - Y^n_k X^n_{in}}{|X^n_{in}| + |X^n_{kn}|}, & \text{if } |X^n_{in}| + |X^n_{kn}| \neq 0; \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}
\tag{12}
$$

Note that if $|X^n_{in}| + |X^n_{kn}| = 0$, then the above formulas produce a row of zeros that corresponds to a zero term, which does not contribute to the objective function. We assume that all such zero rows are removed, and the rest of the rows are renumbered (reindexed) to preserve continual enumeration.

We denote the number of nonzero rows by M_{n-1} and observe that

$$M_{n-1} \leq M_n(M_n - 1)/2.$$

Finally, we take the numbers X^{n-1}_{ij} and Y^{n-1}_i with $i = 1, \ldots, M_{n-1}$ and $j = 1, \ldots, n-1$ to form the matrix and vector

$$\boldsymbol{X}_{n-1} = (X^{n-1}_{ij}), \qquad \boldsymbol{Y}_{n-1} = (Y^{n-1}_i),$$

which completely determine the objective function in the reduced problem. Specifically, the reduced problem for $n = 1$ degenerates into the constant

$$\mu = \|\boldsymbol{Y}_0\|, \tag{13}$$

representing the minimum of the objective function in the initial problem.

5.2. Derivation of Box Constraints

We take the double inequality at (3) and adjust it to write the box constraint for the parameter θ_n in the form

$$
\max_{\substack{1 \leq i \leq M_n \\ X^n_{in} \neq 0}} \left(-\frac{\mu}{|X^n_{in}|} - \sum_{j=1}^{n-1} \frac{X^n_{ij}}{X^n_{in}} \theta_j + \frac{Y^n_i}{X^n_{in}} \right) \leq \theta_n \leq \min_{\substack{1 \leq i \leq M_n \\ X^n_{in} \neq 0}} \left(\frac{\mu}{|X^n_{in}|} - \sum_{j=1}^{n-1} \frac{X^n_{ij}}{X^n_{in}} \theta_j + \frac{Y^n_i}{X^n_{in}} \right),
$$

where μ denotes the minimum value of the objective function in problem (1).

To represent this inequality constraint in vector form, we introduce the following notation. First, for all $i = 1, \ldots, M_n$ such that $X^n_{in} \neq 0$ and all $j = 1, \ldots, n-1$, we define

$$T^n_{ij} = \frac{X^n_{ij}}{X^n_{in}}, \qquad L^n_i = \frac{Y^n_i}{X^n_{in}} - \frac{\mu}{|X^n_{in}|}, \qquad U^n_i = \frac{Y^n_i}{X^n_{in}} + \frac{\mu}{|X^n_{in}|}. \tag{14}$$

We note that all indices i with $X^n_{in} = 0$ are not taken into account when calculating the maximum and minimum in the double inequality, and hence are excluded from the index set $1, \ldots, M_n$. Assuming that the rest of indices are renumbered to preserve continual enumeration, we introduce the matrix and column vectors

$$\boldsymbol{T}_n = (T^n_{ij}), \qquad \boldsymbol{L}_n = (L^n_i), \qquad \boldsymbol{U}_n = (U^n_i).$$

With this matrix-vector notation, we write the box constraint, if $n > 1$, as the double inequality

$$\max(\boldsymbol{L}_n - \boldsymbol{T}_n \boldsymbol{\theta}_{n-1}) \le \theta_n \le \min(\boldsymbol{U}_n - \boldsymbol{T}_n \boldsymbol{\theta}_{n-1}), \tag{15}$$

and, if $n = 1$, as the inequality

$$\max(\boldsymbol{L}_1) \le \theta_1 \le \min(\boldsymbol{U}_1), \tag{16}$$

where max and min symbols are thought of as operators that calculate the maximum and minimum over all elements of corresponding vectors.

5.3. Solution Algorithm

We summarize the above consideration in the form of a computational algorithm to solve problem (1) in a finite number of steps. The algorithm includes two sequential phases: backward elimination and forward determination (substitution) of the unknown parameters.

The backward elimination starts with $n = N$ by setting $M_n = M$ and

$$\boldsymbol{X}_n = \boldsymbol{X}, \qquad \boldsymbol{Y}_n = \boldsymbol{Y}.$$

Furthermore, for each $n = N, \ldots, 1$, the matrix $\boldsymbol{X}_n$ and vector $\boldsymbol{Y}_n$ are used as described above to obtain the values of $\boldsymbol{X}_{n-1}$ and $\boldsymbol{Y}_{n-1}$ if $n > 1$ or the value of $\boldsymbol{Y}_0$ if $n = 1$. As supplementary results, the matrices $\boldsymbol{T}_n$ are also evaluated from the matrices $\boldsymbol{X}_n$.

The backward elimination completes at $n = 1$ by calculating the minimum value of the objective function, given by $\mu = \|\boldsymbol{Y}_0\|$.

The forward determination first uses the obtained minimum μ, matrix $\boldsymbol{X}_1$ and vector $\boldsymbol{Y}_1$ to calculate the vectors $\boldsymbol{L}_1$ and $\boldsymbol{U}_1$ and then evaluate the box constraint for the unknown θ_1 in the form of double inequality at (16). Then, for each $n = 2, \ldots, N$, the vectors $\boldsymbol{L}_n$ and $\boldsymbol{U}_n$ are calculated from μ, $\boldsymbol{X}_n$ and $\boldsymbol{Y}_n$ to represent the box constraints for θ_n as in (15).

Note that the bounds in the box constraint for the parameter θ_1 are explicitly defined by constants, whereas the bounds for each parameter θ_n with $n > 1$ are defined as linear functions of the previous parameters $\theta_1, \ldots, \theta_{n-1}$. Therefore, we can first fix a value to satisfy the box constraint for θ_1 and then substitute this value into the box constraint for θ_2 to obtain explicit bounds given by constants. By repeating such calculations to fix a value for a parameter with explicit bounds and to evaluate bounds for the next parameter, we can successively determine a solution of the problem.

5.4. Computational Complexity

The most computationally intensive and memory demanding component of the algorithm, which determines the overall rate of computational complexity and memory requirement, is the calculation of entries in the matrices $\boldsymbol{X}_n$ and vectors $\boldsymbol{Y}_n$ for all $n = N - 1, N - 2, \ldots, 1$, and vector $\boldsymbol{Y}_0$ by using (12). Though calculating one entry involves a few simple operations, the number of all entries grows very fast as M and N increase.

To derive a rough estimate for the number of entries in all matrices, we first evaluate the number of rows in each matrix. Assuming that the matrix $\boldsymbol{X}_N = \boldsymbol{X}$ has M rows, we see that the number of rows in $\boldsymbol{X}_{N-1}$ is bounded from above by

$$M_{N-1} \le M_N(M_N - 1)/2 = M(M - 1)/2 < M^2/2 = 2(M/2)^2.$$

Recursive application of this estimate yields an upper bound for the number of rows in the matrices $\boldsymbol{X}_{N-l}$ for each $l = 1, \ldots, N - 1$ in the form

$$M_{N-l} \le M_{N-l+1}(M_{N-l+1} - 1)/2 < M_{N-l+1}^2/2 < 2(M/2)^{2^l}.$$

At the last step with $l = N$, we calculate the vector $\boldsymbol{Y}_0$, in which the number of entries is no more than

$$M_0 \le M_1(M_1 - 1)/2 < M_1^2/2 < 2(M/2)^{2^N}.$$

Since we have $n + 1$ columns in the matrix $\boldsymbol{X}_n$ together with the vector $\boldsymbol{Y}_n$, the overall number of the entries in all steps can be estimated as

$$\sum_{l=1}^{N}(N - l + 1)M_{N-l} < \sum_{l=1}^{N} 2(N - l + 1)(M/2)^{2^l}.$$

We denote the upper bound on the right-hand side by $C(N, M)$ and observe that this bound increases polynomially with respect to M, and double exponentially with N. For problems with few parameters, the value of $C(N, M)$ seems to be rather acceptable. Specifically, in the three-parameter case with $N = 3$ and $M = 10,20,50$, we have $C(3, 10) = 783,900$, $C(3, 20) = 200,040,600$, and $C(3, 50) = 305,177,347,500$. A further increase of the number of parameters N results in a rapid rise in the value of $C(N, M)$, as the following examples show: $C(4, 10) = 305,177,347,700$, $C(5, 10) = 610,353,911,500$.

Note that the actual number of entries to calculate is fewer than that given by the bound $C(N, M)$. As it follows from (12), this number can be further reduced if the matrices $\boldsymbol{X}_n$, or at least the initial matrix $\boldsymbol{X}_N = \boldsymbol{X}$, have many zero entries (sparse matrices). It is clear that, if a matrix has zero entries in a column other than the last one, the columns (and related parameters) can be renumbered to put the column with zero entries on the last place where zero entries can reduce computations. As an example of the case with good chances of having sparse matrices $\boldsymbol{X}_n$, one can consider problems where the initial matrix $\boldsymbol{X}$ has entries that take values from the set $\{-1, 0, 1\}$.

Finally, we observe that the calculations by Formula (12) can be performed for different entries quite independently, which offers strong potential for parallel implementation of the procedure on parallel and vector computers to provide more computational and memory resources, and hence to extend applicability to problems of higher dimensions.

6. Software Implementation and Numerical Examples

We conclude with comments on software implementation of the solution procedure and illustrative numerical examples of low dimensions that demonstrate the computational technique involved in the solution. For the sake of illustration, we concentrate on application to problems with exact input data given by integer (rational) numbers to find explicit rational solutions.

To obtain exact solutions, the procedure has been coded for serial symbolic computations in the MATLAB (Release R2020b) environment as a collection of functions that calculate all intermediate matrices and vectors, as well as provide the overall functionality of the algorithm. The numerical experiments were conducted on a custom computer with a 4-core 8-thread Intel Xeon E3-1231 v3 CPU at 3.40GHz and 32GB of DDR3 RAM, running Windows 10 Enterprise 64-bit OS.

Example 1. *Let us take $N = 3$ and $M = 4$ and consider the approximation (regression) problem with*

$$\boldsymbol{X} = \begin{pmatrix} 3 & -1 & 2 \\ -1 & -2 & 2 \\ -2 & 3 & -1 \\ 0 & 2 & -1 \end{pmatrix}, \qquad \boldsymbol{Y} = \begin{pmatrix} 2 \\ 1 \\ -1 \\ 0 \end{pmatrix}.$$

We start with the backward elimination phase by setting $\boldsymbol{X}_3 = \boldsymbol{X}$ and $\boldsymbol{Y}_3 = \boldsymbol{Y}$. We use (12) to calculate the entries in

$$X_2 = \begin{pmatrix} 2 & 1/2 \\ 1/3 & -5/3 \\ -1 & -1 \\ 5/3 & -4/3 \\ 1/3 & -2/3 \\ 1 & -1/2 \end{pmatrix}, \qquad Y_2 = \begin{pmatrix} 1/2 \\ 0 \\ -2/3 \\ 1/3 \\ -1/3 \\ 1/2 \end{pmatrix},$$

and then the entries in

$$X_1 = \begin{pmatrix} -21/13 \\ -1 \\ -21/11 \\ -9/7 \\ -3/2 \\ -3/4 \\ 7/9 \\ 1/7 \\ 9/13 \\ 9/7 \\ 3/5 \\ 1 \\ -1/3 \\ 3/11 \\ 3/7 \end{pmatrix}, \qquad Y_1 = \begin{pmatrix} -5/13 \\ -1/9 \\ -5/11 \\ -1/7 \\ -1/2 \\ -5/12 \\ 5/27 \\ -5/21 \\ 5/13 \\ 11/21 \\ 1/15 \\ 5/9 \\ -1/3 \\ 3/11 \\ 3/7 \end{pmatrix}.$$

Next, we obtain the vector Y_0, which appears to have 105 entries and thus is not shown here to save space. Evaluating the maximum entry of Y_0 as the minimum of the objective function according to (13) yields

$$\mu = 2/7.$$

In parallel with evaluating the entries of X_n and Y_n, we apply (14) to find the entries in

$$T_3 = \begin{pmatrix} -3/2 & 1/2 \\ 1/2 & 1 \\ -2 & 3 \\ 0 & 2 \end{pmatrix}, \qquad T_2 = \begin{pmatrix} -4 \\ 1/5 \\ -1 \\ 5/4 \\ 1/2 \\ 2 \end{pmatrix}.$$

Substitution of the obtained minimum $\mu = 2/7$ yields

$$L_3 = \begin{pmatrix} 6/7 \\ 5/14 \\ 5/7 \\ -2/7 \end{pmatrix}, \qquad U_3 = \begin{pmatrix} 8/7 \\ 9/14 \\ 9/7 \\ 2/7 \end{pmatrix}, \qquad L_2 = \begin{pmatrix} 3/7 \\ -6/35 \\ 8/21 \\ -13/28 \\ 1/14 \\ -11/7 \end{pmatrix}, \qquad U_2 = \begin{pmatrix} 11/7 \\ 6/35 \\ 20/21 \\ -1/28 \\ 13/14 \\ -3/7 \end{pmatrix}.$$

Finally, we calculate the vectors

$$L_1 = \begin{pmatrix} 3/49 \\ -11/63 \\ 13/147 \\ -1/9 \\ 1/7 \\ 11/63 \\ -19/147 \\ -11/3 \\ 1/7 \\ 5/27 \\ -23/63 \\ 17/63 \\ 1/7 \\ -1/21 \\ 1/3 \end{pmatrix}, \qquad U_1 = \begin{pmatrix} 61/147 \\ 25/63 \\ 19/49 \\ 1/3 \\ 11/21 \\ 59/63 \\ 89/147 \\ 1/3 \\ 61/63 \\ 17/27 \\ 37/63 \\ 53/63 \\ 13/7 \\ 43/21 \\ 5/3 \end{pmatrix}.$$

The forward determination (substitution) phase of the procedure involves application of (16) and then (15) to obtain the unique solution

$$\theta_1 = 1/3, \qquad \theta_2 = 5/21, \qquad \theta_3 = 16/21.$$

The actual number of entries in the matrices X_n and vectors Y_n to calculate in the problem is 153, which is fewer than the upper bound given by $C(3,4) = 600$. The computer time to obtain the exact solution by symbolic computation averages 2.3 s.

Example 2. *Consider the problem with $N = 3$ and $M = 10$, where the input data are given by*

$$X = \begin{pmatrix} 3 & -1 & 2 \\ -1 & -2 & 2 \\ -2 & 3 & -1 \\ 0 & 2 & -1 \\ 1 & 2 & -1 \\ 3 & 1 & 0 \\ 1 & 1 & -1 \\ -1 & -1 & 2 \\ 0 & 3 & 1 \\ 2 & 1 & 0 \end{pmatrix}, \qquad Y = \begin{pmatrix} 2 \\ 1 \\ -1 \\ 0 \\ 1 \\ -1 \\ 2 \\ 0 \\ 1 \\ 3 \end{pmatrix}.$$

The application of the procedure yields the unique solution

$$\theta_1 = 1/9, \qquad \theta_2 = 13/18, \qquad \theta_3 = 8/9.$$

The matrices X_n and vectors Y_n involved in calculations have $444,280$ entries (while the corresponding upper bound is $C(3,10) = 783{,}900$). The symbolic computations take about 52 min of computer time.

7. Conclusions

A direct computational technique has been proposed to solve discrete linear Chebyshev approximation problems, which find wide application in various areas, including the least maximum absolute deviation regression in statistics. First, we have shown that the problem under consideration can be reduced by eliminating an unknown parameter to a problem with less number of unknowns and a box constraint for the parameter eliminated. This result was used to obtain direct solutions to linear regression problems with one and two parameters.

To solve approximation problems of arbitrary dimension, we have developed a direct solution procedure that consists of two parts: backward elimination and forward substitution of the unknown parameters. The direct solution is of particular interest in the problems when analytical solutions are desired, whereas the use of iterative algorithms may be inappropriate or inadequate. We have estimated the computational complexity of the procedure, discussed its MATLAB implementation intended to provide exact solutions by symbolic computations, and presented numerical examples.

Possible lines of further investigation can include modification and improvement of the algorithm to reduce computational complexity in solving problems in both exact (rational) and inexact (floating point) form. The development of parallel implementations of the algorithm to speed up calculations is also of interest.

Conflicts of Interest: The author declares no conflict of interest.

References

1. de Laplace, P.S. *Mécanique Céleste*; Hillard, Gray, Littlè, and Wilkins: Boston, MA, USA, 1832; Volume 2. (Engl. Transl. with Comment. by N. Bowditch)
2. Harter, H.L. The method of least squares and some alternatives: Part I. *Int. Stat. Rev.* **1974**, *42*, 147–174. [CrossRef]
3. Steffens, K.G. *The History of Approximation Theory*; Birkhäuser: Boston, MA, USA, 2006. [CrossRef]
4. Tewarson, R.P. On minimax solutions of linear equations. *Comput. J.* **1972**, *15*, 277–279. [CrossRef]
5. Appa, G.; Smith, C. On L_1 and Chebyshev estimation. *Math. Program.* **1973**, *5*, 73–87. [CrossRef]
6. Pĭnar, M.Ç. Overdetermined systems of linear equations. In *Encyclopedia of Optimization*, 2nd ed.; Floudas, C.A., Pardalos, P.M., Eds.; Springer: Boston, MA, USA, 2009; pp. 2893–2896. [CrossRef]
7. Rabinowitz, P. Applications of linear programming to numerical analysis. *SIAM Rev.* **1968**, *10*, 121–159. [CrossRef]
8. Hand, M.L. Aspects of Linear Regression Estimation Under the Criterion of Minimizing the Maximum Absolute Residual. Ph.D. Thesis, Iowa State University, Ames, IA, USA, 1978.
9. Kennedy, W.J.; Gentle, J.E. *Statistical Computing*; Statistics: Textbooks and Monographs; Dekker: New York, NY, USA, 1980; Volume 33.
10. James, F. Fitting tracks in wire chambers using the Chebyshev norm instead of least squares. *Nucl. Instr. Meth.* **1983**, *211*, 145–152. [CrossRef]
11. Bertsch, G.F.; Sabbey, B.; Uusnäkki, M. Fitting theories of nuclear binding energies. *Phys. Rev. C* **2005**, *71*, 054311. [CrossRef]
12. Mäkilä, P.M. Robust identification and Galois sequences. *Internat. J. Control* **1991**, *54*, 1189–1200. [CrossRef]
13. Akçay, H.; Hjalmarsson, H.; Ljung, L. On the choice of norms in system identification. *IEEE Trans. Automat. Control* **1996**, *41*, 1367–1372. [CrossRef]
14. Schölkopf, B.; Smola, A.J. *Learning with Kernels*; Adaptive Computation and Machine Learning; MIT Press: Cambridge, MA, USA, 2002; p. 626.
15. Bartoszuk, M.; Beliakov, G.; Gagolewski, M.; James, S. Fitting aggregation functions to data: Part I—Linearization and regularization. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems*; Carvalho, J.P., Lesot, M.J., Kaymak, U., Vieira, S., Bouchon-Meunier, B., Yager, R.R., Eds.; Springer: Cham, Switzerland, 2016; Volume 611, pp. 767–779. [CrossRef]
16. Jaschke, S.R. Arbitrage bounds for the term structure of interest rates. *Finance Stoch.* **1997**, *2*, 29–40. [CrossRef]
17. Harter, H.L. The method of least squares and some alternatives: Part III. *Int. Stat. Rev.* **1975**, *43*, 1–44. [CrossRef]
18. Harter, H.L. The method of least squares and some alternatives: Part IV. *Int. Stat. Rev.* **1975**, *43*, 125–190. [CrossRef]

19. Späth, H. *Mathematical Algorithms for Linear Regression*; Computer Science and Scientific Computing; Academic Press, Inc.: San Diego, CA, USA, 1992; p. 338.

20. Wagner, H.M. Linear programming techniques for regression analysis. *J. Am. Statist. Assoc.* **1959**, *54*, 206–212. [CrossRef]

21. Stiefel, E. Note on Jordan elimination, linear programming and Tchebycheff approximation. *Numer. Math.* **1960**, *2*, 1–17. [CrossRef]

22. Watson, G.A. On the best linear one-sided Chebyshev approximation. *J. Approx. Theory* **1973**, *7*, 48–58. [CrossRef]

23. Sposito, V.A. Minimizing the maximum absolute deviation. *ACM SIGMAP Bull.* **1976**, 51–53. [CrossRef]

24. Armstrong, R.D.; Kung, D.S. Algorithm AS 135: Min-max estimates for a linear multiple regression problem. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **1979**, *28*, 93–100. [CrossRef]

25. Kim, B.Y. Algorithm for the constrained Chebyshev estimation in linear regression. *Commun. Stat. Appl. Methods* **2000**, *7*, 47–54.

26. Boyd, S.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press: Cambridge, UK, 2004.

27. Castillo, E.; Mínguez, R.; Castillo, C.; Cofiño, A.S. Dealing with the multiplicity of solutions of the ℓ_1 and ℓ_∞ regression models. *Eur. J. Oper. Res.* **2008**, *188*, 460–484. [CrossRef]

28. Ene, A.; Vladu, A. Improved convergence for ℓ_1 and ℓ_∞ regression via iteratively reweighted least squares. In Proceedings of the 36th International Conference on Machine Learning Research, Long Beach, CA, USA, 10–15 June 2019; Chaudhuri, K., Salakhutdinov, R., Eds.; Volume 97, pp. 1794–1801.

29. Krivulin, N. Complete solution of a constrained tropical optimization problem with application to location analysis. In *Relational and Algebraic Methods in Computer Science*; Lecture Notes in Computer Science; Höfner, P., Jipsen, P., Kahl, W., Müller, M.E., Eds.; Springer: Cham, Switzerland, 2014; Volume 8428, pp. 362–378. [CrossRef]

30. Krivulin, N. Algebraic solution of weighted minimax single-facility constrained location problems. In *Relational and Algebraic Methods in Computer Science*; Lecture Notes in Computer Science; Desharnais, J., Guttmann, W., Joosten, S., Eds.; Springer: Cham, Switzerland, 2018; Volume 11194, pp. 317–332. [CrossRef]

31. Krivulin, N. Algebraic solution of minimax single-facility constrained location problems with Chebyshev and rectilinear distances. *J. Log. Algebr. Methods Program.* **2020**, *115*, 100578. [CrossRef]