mathematics

MDPI

Article

Systems of Linear Equations with Non-Negativity Constraints: Hyper-Rectangle Cover Theory and Its Applications

Xiaoxuan Chu, Kon Max Wong *, Jun Chen and Jiankang Zhang †

Department of Electrical and Computer Engineering, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4L8, Canada; chux7@mcmaster.ca (X.C.); chenjun@mcmaster.ca (J.C.); jkzhang@mcmaster.ca (J.Z.)
* Correspondence: wongkm@mcmaster.ca
† Deceased.

Abstract: In this paper, a novel hyper-rectangle cover theory is developed. Two important concepts, the *cover order* and the *cover length*, are introduced. We construct a specific échelon form of the matrix in the same manner as that employed to determine the rank of the matrix to obtain the cover order of any given matrix. Using the properties of the cover order, we obtain the necessary and sufficient conditions for the existence and uniqueness of the solutions for linear equations system with *non-negativity constraints* on variables for both *homogeneous* and *nonhomogeneous* cases. In addition, we apply the cover theory to analyze some typical problems in linear algebra and optimization with non-negativity constraints on variables, including *linear programming* (LP) problems and *non-negative least squares* (NNLS) problems. For LP problems, the three possible behaviours of the solutions are studied through cover theory. On the other hand, we develop a method to obtain the cover length of the covered variable. In this process, we discover the relationship between the cover length determination problem and the NNLS problem. This enables us to obtain an analytical optimal value for the NNLS problem.

Keywords: hyper-rectangle cover; cover order; cover length; system of linear equations; non-negativity constraints; non-negative least squares; linear programming

MSC: 15A06; 90C05; 93E24

Citation: Chu, X.; Wong, K.M.; Chen, J.; Zhang, J. Systems of Linear Equations with Non-Negativity Constraints: Hyper-Rectangle Cover Theory and Its Applications. *Mathematics* **2023**, *11*, 2338. https://doi.org/10.3390/math11102338

Academic Editor: Andrey Amosov

Received: 6 March 2023
Revised: 8 May 2023
Accepted: 12 May 2023
Published: 17 May 2023

1. Introduction

The problems with non-negativity constraints on variables play a prominent role in engineering, physics, chemistry, computer science, and economics. These problems with non-negative constraints often appear as (1) finding solutions for systems of linear equations, (2) solving LP problems, and (3) finding solutions for NNLS problems [1].

The analysis of systems of linear equations is a fundamental part of linear algebra, and forms the core of mathematical modelling of many different branches of science and engineering such as, to name but a few, electric circuits, communications, radars, optics, controls, etc. [2–6]. Lately, it has also been used to model the outbreak of COVID-19 and calcium diffusion [7–9]. Thus, methods for finding the solutions of linear equations system also play an important role in various applications [10]. Some mature methods have been developed to analyze and solve systems of linear equations without non-negativity constraints on the variables [11–13]. However, with non-negativity constraints added to the variables, the analysis of the solutions to the linear equations becomes harder [14]. For such problems, the classical analysis for the existence of non-negative solutions is mainly based on Farkas' lemma [15]. In terms of uniqueness, there is no direct characterization in the general case. It is also noted that the analysis of non-negative solutions to the system of homogeneous or nonhomogeneous linear equations is mostly concerned with investigating other associated problems rather than addressing the problem in a direct way. Thus, a new approach is needed for the analysis of systems of linear equations with non-negativity constraints on the variables.

LP problems arise in many applications [16]. Many problems can be reformulated as linear programs both in theory and in practice so that fast algorithms can be applied [17–20]. Dantzig developed the simplex method in 1947, which was the first efficient method for solving LP problems and has been widely accepted as a computational tool [21–23]. Geometrically, the procedure of the simplex method involves moving one feasible solution to another and, for each step, the value of the objective function improves. This continues until the optimum objective is reached. It would thus be desirable if we can determine the optimal solution directly rather than moving through the feasible solutions for matrices with some specific structures. In this paper, we propose a new systematic procedure for solving the LP problem by applying cover theory using a transformed objective function.

The problem of NNLS is a type of least squares problem with non-negativity constraints on variables, which arises in applications throughout science and engineering. Various methods have been proposed to solve this kind of problems, and they can normally be divided into three classes: active-set methods [24–26], iterative methods [27,28], and other methods [29–32]. The first technique to solve the NNLS problem was proposed by Lawson and Hanson in [33], which is a typical active-set method, and the corresponding algorithm is named lsqnonneg in Matlab. This commonly used algorithm always converges and terminates in finite steps; however, there is no upper limit on the possible number of iterations that the algorithm might need to reach the point of the optimum value. In contrast to active-set methods, iterative methods enable one to incorporate multiple active constraints at each iteration. Since most existing algorithms for solving the NNLS problem are based on numerical analysis, we are motivated to propose a method to solve it from the matrix perspective by applying the techniques we developed in cover theory. More specifically, we solve the problem by investigating the structure of the matrix itself so as to obtain the analytical optimal value of NNLS problem.

Overview of the Paper

In this paper, we establish the novel hyper-rectangle cover theory for which we obtain the necessary and sufficient conditions that guarantee the existence and the uniqueness of the solution for a system of linear equations with non-negativity constraints on variables. A specific échelon form of the matrix is introduced and based on this form, and an efficient method is developed to determine the cover order for any given matrix. Moreover, we investigate in detail the structures of the échelon form in various cases leading to the development of feasible solutions for the system of linear equations with non-negativity constraints on variables. Parallel investigations are carried out for the LP problems. Based on the échelon form and the corresponding results on the system of linear equations with non-negative constraints, we also analyze the various possibilities of the solution for an LP problem. Finally, we develop a method to determine the cover length of the covered variables, establishing their strong relationship with the optimal objective value of NNLS problems. Based on this relationship, a new method is derived to obtain the analytical optimal value of NNLS problems.

Notation: Most notations used throughout this paper are standard: column vectors and matrices are denoted by boldface lowercase and uppercase characters, respectively; the matrix transpose is denoted by $(\cdot)^T$; $\mathbb{R}_+^N$ denotes the set of all the $N \times 1$ vectors with all entries being non-negative.

2. Concept of Hyper-Rectangle Cover

In this section, we formally give the definition of hyper-rectangle cover [34,35].

Definition 1. *Given a matrix* $\mathbf{A} \in \mathbb{R}^{M \times N}$ *and* $\mathbf{x} \in \mathbb{R}_+^N$, x_n *is the n-th element in* $\mathbf{x}$ *for index* $n \in \{1, 2, \cdots, N\}$. *Let* $c_n(x_n)$ *denote the smallest number* c_n *such that*

$$\left\{ \mathbf{x} \in \mathbb{R}_+^N : \ \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} \leq 1 \right\} \subseteq \left\{ \mathbf{x} \in \mathbb{R}_+^N : \ x_n \leq c_n \right\} \tag{1}$$

We say that x_n *is a covered variable if* $c_n(x_n)$ *is finite and we refer to* $c_n(x_n)$ *as the cover length of the covered variable* x_n. *The cover order of* $\mathbf{A}$, *denoted by* $R_c(\mathbf{A})$, *is the number of indices* $n \in \{1, 2, \cdots, N\}$ *for which* $c_n(x_n) < \infty$. *We say that* $\mathbf{A}$ *has full cover if* $R_c(\mathbf{A}) = N$ *and has zero cover if* $R_c(\mathbf{A}) = 0$.

The following nontrivial examples may serve as illustrations of the definitions of cover order and cover length.

Example 1. *Suppose that* $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, *and we have* $\mathbf{A}^T \mathbf{A} = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$. *For this matrix,* $\{\mathbf{x} \in \mathbb{R}_+^2 : \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} \leq \tau^2\}$ *is an ellipse in the whole plane, which is shown in Figure 1, and the part which is located in the non-negative domain is fully covered by a rectangle. Thus,* x_1 *and* x_2 *are both covered in this example.*

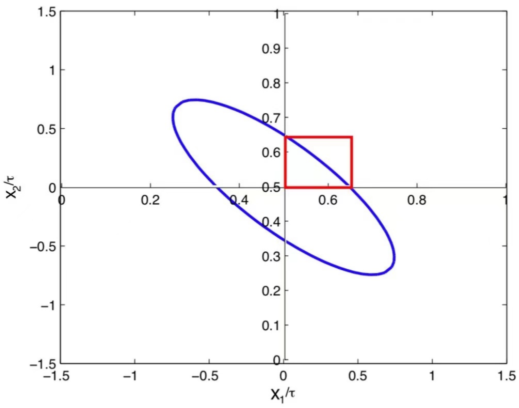

Figure 1. Example of a 2×2 full-cover matrix.

Example 2. *Consider* $\mathbf{A} = \begin{pmatrix} 1 & -1 \end{pmatrix}$ *and* $\mathbf{A}^T \mathbf{A} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$; *for this case, the feasible set determined by* $\mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} = (x_1 - x_2)^2 \leq \tau^2$ *is shown in Figure 2 and is open and unbounded with respect to both* x_1 *and* x_2. *Hence,* $R_c(\mathbf{A}) = 0$.

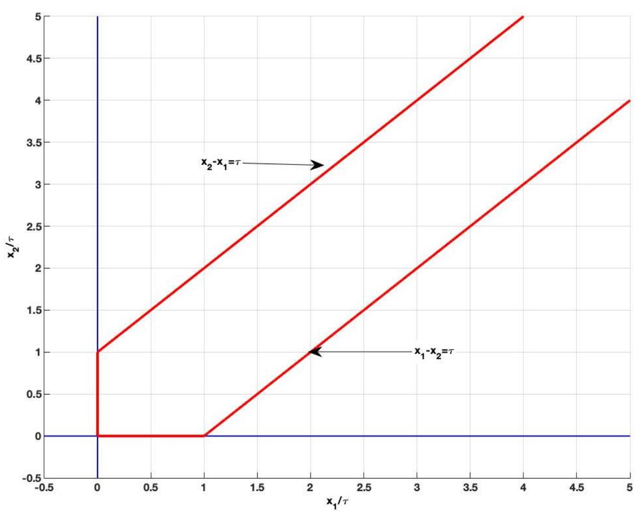

Figure 2. Example of a zero-cover matrix.

From the above examples, we can find that the cover order $R_c(\mathbf{A})$ and the cover length $c_i(x_i)$ represent the maximal dimension and minimal side lengths of the hyper-rectangle that covers $\{\mathbf{x} : \mathbf{x} \in \mathbb{R}_+^N, \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2\}$ respectively.

3. Systems of Linear Equations with Non-Negativity Constraints on Solutions

3.1. Homogeneous Systems of Linear Equations

In this section, we present an important result in cover theory, allowing us to determine if a column vector in $\mathbf{A}$ or the corresponding variable x_i in $\mathbf{Ax}$ is covered or not. Furthermore, this result provides us with a method of investigating the non-negative solution to a system of linear equations. Let us first introduce the definition of a convex cone [36].

Definition 2. *A set $\mathcal{C} \subseteq \mathbb{R}^M$ is a convex cone if $\alpha\mathbf{x} + \beta\mathbf{y} \in \mathcal{C}$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{C}$ and $\alpha, \beta > 0$.*

A cone $\mathcal{C}$ is *polyhedral* if it is the conic combination of finitely many vectors, i.e., there is a set of vectors $\{\mathbf{a}_1, \cdots, \mathbf{a}_N\}$, so that $\mathcal{C} = \{\mathbf{a}_1u_1 + \mathbf{a}_2u_2 + \cdots \mathbf{a}_Nu_N | \mathbf{a}_i \in \mathbb{R}^M, u_i \in \mathbb{R}^+\}$. The polyhedral cone $\mathcal{C}$ is a closed convex cone.

With the above definition, we are now able to obtain the following result:

Theorem 1. *Let $\mathbf{A}$ be an $M \times N$ real matrix. Then, the i-th column of $\mathbf{A}$, or the i-th variable x_i associated with the i-th column vector $\mathbf{a}_i$ in $\mathbf{Ax}$, is covered if and only if $\mathbf{Ax} \neq \mathbf{0}$ for any $\mathbf{x} \in \mathbb{R}_+^N$ with $x_i > 0$.*

Proof. *Necessity*: Here, by assuming that x_i is covered in $\mathbf{Ax}$, we need to show that for any $\mathbf{x} \in \mathbb{R}_+^N$ with $x_i > 0$, we have $\mathbf{Ax} \neq \mathbf{0}$. Suppose that this statement was not true. Then, there exists $\mathbf{x}_0 \in \mathbb{R}_+^N$ with $x_{0,i} > 0$ such that $\mathbf{Ax}_0 = \mathbf{0}$, where $x_{0,i}$ is the i-th element in $\mathbf{x}_0$. As a consequence, for any positive number $p > 0$, we would also have $\mathbf{A}(p\mathbf{x}_0) = \mathbf{0}$, implying that $px_{0,i}$ is not bounded if for any given $\tau > 0$, $0 = (p\mathbf{x}_0)^T\mathbf{A}^T\mathbf{A}(p\mathbf{x}_0) \leq \tau$. This contradicts the assumption that x_i is covered in $\mathbf{Ax}$. Therefore, the necessary condition is true.

Sufficiency: For $x_i > 0$, the quadratic form $\mathbf{x}^T\mathbf{A}^T\mathbf{Ax}$ can be rewritten as: $\mathbf{x}^T\mathbf{A}^T\mathbf{Ax} = \|\bar{\mathbf{A}}_i\bar{\mathbf{x}}_i + \mathbf{a}_ix_i\|^2 = x_i^2\|\bar{\mathbf{A}}_i\mathbf{u} + \mathbf{a}_i\|^2$, where $\mathbf{u} = \frac{\bar{\mathbf{x}}_i}{x_i}$, $\bar{\mathbf{A}}_i$ is the $M \times (N-1)$ sub-matrix formed by deleting the i-th column from $\mathbf{A}$ and $\mathbf{u} \geq \mathbf{0}$. Consider the set $\left\{\bar{\mathbf{A}}_i\mathbf{u} : \mathbf{u} \in \mathbb{R}_+^{N-1}\right\}$. It is a closed convex cone according to Definition 2, and the function $\|\bar{\mathbf{A}}_i\mathbf{u} + \mathbf{a}_i\|^2$ is convex; thus, the minimum of $\|\bar{\mathbf{A}}_i\mathbf{u} + \mathbf{a}_i\|^2$ exists, i.e., there exists a $\mathbf{u}_0 \geq \mathbf{0}$ such that $\|\bar{\mathbf{A}}_i\mathbf{u}_0 + \mathbf{a}_i\|^2 \leq \|\bar{\mathbf{A}}_i\mathbf{u} + \mathbf{a}_i\|^2$, for any $\mathbf{u} \in \mathbb{R}_+^{N-1}$. In fact, $\|\bar{\mathbf{A}}_i\mathbf{u}_0 + \mathbf{a}_i\| \neq 0$. Otherwise, if $x_{0,i} = 1, x_{0,k} = u_{0,k}$, for $k = 1, 2, \cdots, i-1$, and $x_{0,k} = u_{0,k-1}$, for $k = i+1, \cdots, N$, then, we have $\bar{\mathbf{A}}_i\mathbf{u}_0 + \mathbf{a}_i = \mathbf{Ax}_0 = \mathbf{0}$, which contradicts the assumption. Now for any given positive real value τ, if we let $\mathbf{x}^T\mathbf{A}^T\mathbf{Ax} \leq \tau^2$, then, we have $x_i^2\|\bar{\mathbf{A}}_i\mathbf{u}_0 + \mathbf{a}_i\|^2 \leq \mathbf{x}^T\mathbf{A}^T\mathbf{Ax} \leq \tau^2$. Hence, we obtain $0 < x_i < \frac{\tau}{\|\bar{\mathbf{A}}_i\mathbf{u}_0 + \mathbf{a}_i\|}$, i.e., x_i is covered in $\mathbf{Ax}$.

Thus, the proof of Theorem 1 is complete. □

From Theorem 1, the following results can be obtained.

Corollary 1. *Let $\mathbf{A}$ be an $M \times N$ real matrix, and let $\bar{\mathbf{A}}_j$ be the $M \times (N-1)$ sub-matrix formed by deleting the j-th column from $\mathbf{A}$. Then, the following statements are true:*

1. *A system of homogeneous linear equations, $\mathbf{Ax} = \mathbf{0}$, has a nonzero solution in $\mathbb{R}_+^N$ if and only if $\mathbf{A}$ does not have full cover.*
2. *Let the j-th column of $\mathbf{A}$ be covered. Then, any column of $\bar{\mathbf{A}}_j$ is covered in $\bar{\mathbf{A}}_j$ if and only if it, as a column of $\mathbf{A}$, is also covered in $\mathbf{A}$.*
3. *If the i-th column of $\mathbf{A}$ is covered, then it is also covered in $\bar{\mathbf{A}}_j$ for $j \neq i$.*
4. *A full column rank matrix $\mathbf{A}$ always has full cover.*

For a homogeneous system of linear equations with non-negative constraints on solutions, it is important to determine the necessary and sufficient condition which guarantees the existence of nonzero solutions. The direct determination of whether the system has nonzero solution is not simple [14,37]. Here, Theorem 1 makes a statement paralleled to the first statement of Corollary 1, providing us with the condition for the existence of nonzero solutions for the system.

3.2. Nonhomegeneous Systems of Linear Equations with Non-Negativity Constraints on Solutions

Nonhomogeneous systems of linear equations with non-negativity constraints on solutions are frequently encountered in the field of signal and image processing, multispectral data handling, fibre optics, etc. [38–40]. The classical way for determining the existence of non-negative solution is based on Farkas' lemma [15]. According to Farkas' lemma, given a problem of linear equations with non-negativity constraints on the variables, there exists another problem associated with it such that the original problem has a solution in the required domain if and only if the associated problem has no solution. Thus, this lemma provides an indirect way to check the existence of non-negative solutions to a nonhomogeneous system of linear equations.

In the following, based on the cover theory, we will derive the direct necessary and sufficient conditions for the existence and the uniqueness of non-negative solutions of the aforementioned system.

Existence of non-negative solutions:

Theorem 2. *Let* $\mathbf{A} \in \mathbb{R}^{M\times N}$ *and* $\mathbf{b} \in \mathbb{R}^{M}$*. Then, there exists an* $\mathbf{x} \in \mathbb{R}_+^N$ *such that* $\mathbf{Ax} = \mathbf{b}$ *if and only if the cover order of the augmented matrix* $\tilde{\mathbf{A}} = (\mathbf{A}, -\mathbf{b})$ *is less than or equal to that of* $\mathbf{A}$*.*

Proof. Let us rewrite the linear equations $\mathbf{Ax} = \mathbf{b}$ as $\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \mathbf{0}$, where $\tilde{\mathbf{A}} = (\mathbf{A}, -\mathbf{b})$ and $\tilde{\mathbf{x}} = (\mathbf{x}, \tilde{x}_{N+1})$. First we prove the sufficiency: Under the assumption $R_c(\tilde{\mathbf{A}}) \leq R_c(\mathbf{A})$, by Statement 2 of Corollary 1, we can claim that $\tilde{x}_{N+1}$ is not covered in $\tilde{\mathbf{A}}\tilde{\mathbf{x}}$. Since $\tilde{\mathbf{A}}$ does not have full cover, then by Theorem 1, there exists an $\tilde{\mathbf{x}}_0 = (\mathbf{x}_0, \tilde{x}_{0,N+1}) \in \mathbb{R}_+^{N+1}$ with $\tilde{x}_{0,N+1} > 0$, where $\tilde{x}_{0,N+1}$ is the $(N+1)$-th element of $\tilde{\mathbf{x}}_0$, such that $\tilde{\mathbf{A}}\tilde{\mathbf{x}}_0 = \mathbf{0}$. Hence, we have $\mathbf{Ax}_0 = \mathbf{b}\tilde{x}_{0,N+1}$, implying that $\mathbf{x}_0/\tilde{x}_{0,N+1}$ is a solution of $\mathbf{Ax} = \mathbf{b}$. Therefore, the sufficient condition is true.

To prove the necessary condition, we assume that the system of linear equations $\mathbf{Ax} = \mathbf{b}$ has a solution in $\mathbb{R}_+^N$, i.e., there exists an $\mathbf{x}_0 \in \mathbb{R}_+^N$ such that $\mathbf{Ax}_0 = \mathbf{b}$. Then, by Theorem 1, the $(N+1)$-th column vector of $\tilde{\mathbf{A}}$ is not covered. In addition, by Statement 3) of Corollary 1, we know that if $\tilde{\mathbf{a}}_i$, the i-th column vector of $\tilde{\mathbf{A}}$, $i \in \{1, \cdots, N\}$, is covered in $\tilde{\mathbf{A}}$, then $\tilde{\mathbf{a}}_i$, being a column vector of $\mathbf{A}$, is also covered in $\mathbf{A}$. Therefore, the inequality $R_c(\tilde{\mathbf{A}}) \leq R_c(\mathbf{A})$ holds. This completes the proof of Theorem 2. □

Theorem 2 provides us with the regularization condition for non-negative solutions for a systems of linear algebraic equations. It is interesting to point out that in some application of systems of ordinary differential equations, there are parallel regularizations that provide non-negative solutions [41,42].

Uniqueness of non-negative solutions:

With the definition of convex cone in Definition 2, let us first introduce Carathéordory's theorem [15]:

Lemma 1 (Carathéordory's theorem)**.** *Let* $\mathcal{S} \subseteq \mathbb{R}^M$ *be a finite set, and let* $\mathbf{y} \in \mathbb{R}^M$*. If* $\mathbf{y} \in \text{cone}\,\mathcal{S}$*, then there exists a linearly independent set* $\mathcal{T}$ *such that* $\mathbf{y} \in \text{cone}\,\mathcal{T}$*.*

By applying Carathéordory's theorem, we are able to show the necessary and sufficient condition for the *uniqueness* of non-negative solution to a nonhomogeneous system of linear equations. This is stated in the following theorem:

Theorem 3. *Let* $\mathbf{A} \in \mathbb{R}^{M\times N}$ *and* $\mathbf{b} \in \mathbb{R}^M$*. Then,* $\mathbf{Ax} = \mathbf{b}$ *has a unique solution* $\mathbf{x}$ *in* $\mathbb{R}_+^N$*, if and only if* $R_c(\tilde{\mathbf{A}}) \leq R_c(\mathbf{A})$ *and* $R_c(\tilde{\mathbf{A}}) + R_r(\bar{\mathbf{A}}) = N$*, where* $R_r(\bar{\mathbf{A}})$ *is the rank of* $\bar{\mathbf{A}}$*, and* $\bar{\mathbf{A}} = \{\mathbf{a}_i\}_{i\in\bar{\mathcal{N}}}$*, with* $\bar{\mathcal{N}}$ *being a set consisting of all the column indices of* $\mathbf{A}$ *not covered in* $\tilde{\mathbf{A}}$*.*

Proof. *Sufficiency*: Let $\bar{R}_c(\tilde{\mathbf{A}})$ be the number of uncovered column vectors in $\tilde{\mathbf{A}}$. Thus, we have $R_c(\tilde{\mathbf{A}}) + \bar{R}_c(\tilde{\mathbf{A}}) = N + 1$, and under the assumption $R_c(\tilde{\mathbf{A}}) + R_r(\bar{\mathbf{A}}) = N$, we have $\bar{R}_c(\tilde{\mathbf{A}}) = R_r(\bar{\mathbf{A}}) + 1$. In addition, since $R_c(\tilde{\mathbf{A}}) \leq R_c(\mathbf{A})$, we can obtain $\bar{R}_c(\tilde{\mathbf{A}}) = |\bar{\mathcal{N}}| + 1$ and, as a result, $R_r(\bar{\mathbf{A}}) = |\bar{\mathcal{N}}|$. Therefore, all the column vectors $\{\mathbf{a}_i\}_{i\in\bar{\mathcal{N}}}$ are linearly independent in $\mathbb{R}^M$. As a consequence, $\mathbf{A}\mathbf{x} = \mathbf{b}$ has a unique solution in $\mathbb{R}^N_+$.

Necessity: According to the property of $R_c(\tilde{\mathbf{A}})$ and the definition of $\bar{\mathbf{A}}$, we have $R_c(\tilde{\mathbf{A}}) + R_r(\bar{\mathbf{A}}) \leq N$. Considering $R_c(\tilde{\mathbf{A}}) + R_r(\bar{\mathbf{A}}) < N$, in this case, we have $R_r(\bar{\mathbf{A}}) < |\bar{\mathcal{N}}|$. Then, for all $i \in \bar{\mathcal{N}}$, there are $\mu_i \in \mathbb{R}$ (not all 0) such that $\sum_{i\in\bar{\mathcal{N}}} \mu_i \mathbf{a}_i = \mathbf{0}$ and there are $\lambda_i \geq 0$ such that $\sum_{i\in\bar{\mathcal{N}}} \lambda_i \mathbf{a}_i = \mathbf{b}$. In addition, there is a real number α such that $\lambda_i + \alpha\mu_i > 0$ for all $i \in \bar{\mathcal{N}}$. Let us assume that $\mathbf{x}_0 \in \mathbb{R}^N_+$ is the unique solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$, where $x_{0,i} = \lambda_i + \alpha\mu_i$, for $i \in \bar{\mathcal{N}}$ and $x_{0,i} = 0$, for $i \in \{1, 2, \cdots, N\} \setminus \bar{\mathcal{N}}$, then we have $\mathbf{A}\mathbf{x}_0 = \mathbf{b}$. As Carathéordory's theorem states, we are able to find a linearly independent subset $\{\mathbf{a}_i\}_{i\in\tilde{\mathcal{N}}}$ of $\{\mathbf{a}_i\}_{i\in\bar{\mathcal{N}}}$. Let $\tilde{\mathcal{N}}$ be the set consisting of the linearly independent column indices of $\{\mathbf{a}_i\}_{i\in\bar{\mathcal{N}}}$, and $|\tilde{\mathcal{N}}| < |\bar{\mathcal{N}}|$. Then there is another solution $\mathbf{x}_1 \in \mathbb{R}^N_+$ such that $\mathbf{A}\mathbf{x}_1 = \mathbf{b}$ and $x_{1,i} = 0$, for $i \in \{1, 2, \cdots, N\} \setminus \tilde{\mathcal{N}}$, where the number of zero element in $\mathbf{x}_1$ is larger than that in $\mathbf{x}_0$. This contradicts the assumption that $\mathbf{x}_0$ is the unique solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$. Therefore, $R_c(\tilde{\mathbf{A}}) + R_r(\bar{\mathbf{A}}) = N$ must be satisfied to guarantee the uniqueness of the solution. Thus, the proof of Theorem 3 is complete. □

4. Cover Order

In this section, we develop a specific échelon form of a matrix that can be used to determine the cover order of any given matrix. Some special properties of cover order are also explored.

4.1. Cover Order Determination

Let $\mathbb{R}^N_{++}$ denote the set of $N \times 1$ vectors with all entries being positive. The vector $\mathbf{x}$, of which the elements are all positive, is called a *positive* vector. Similarly, let $\mathbb{R}^N_+$, $\mathbb{R}^N_{--}$, and $\mathbb{R}^N_-$, respectively, denote a *non-negative*, a *negative*, and a *nonpositive* set of $N \times 1$ vectors. We first present some related results in the following lemmas which are useful for our derivation of the cover order of a real matrix.

Lemma 2 ([13])**.** *Let $\mathbb{S}$ be a subspace of $\mathbb{R}^N$ and $\mathbb{S}^\perp$ be the orthogonal complementary subspace of $\mathbb{S}$. Then,*

1. $\mathbb{S} \cap \mathbb{R}^N_+ = \emptyset$ *if and only if* $\mathbb{S}^\perp \cap \mathbb{R}^N_{++} \neq \emptyset$.
2. $\mathbb{S} \cap \mathbb{R}^N_{++} = \emptyset$ *if and only if* $\mathbb{S}^\perp \cap \mathbb{R}^N_+ \neq \emptyset$.

Denoting the row space of $\mathbf{A}$ by $\mathbb{S}_\mathbf{A}$ and the orthogonal complement to this row space by $\mathbb{S}^\perp_\mathbf{A}$, and using Lemma 2, we have the following [34]:

Lemma 3. *Let $\mathbb{R}^N_+(K)$ denote the set of all the non-negative vectors with K positive entries. Specifically, $\mathbb{R}^N_+(0)$ denotes the set $\{\mathbf{0}_{N\times 1}\}$. For any $\mathbf{A} \in \mathbb{R}^{M\times N}$, the cover order of $\mathbf{A}$ is equal to $\max_{\mathbb{S}_\mathbf{A} \cap \mathbb{R}^N_+(K) \neq \emptyset} K$.*

Lemma 3 shows us the necessary and sufficient condition for determining the cover order of a matrix. Now, we show an important property of cover order in the following:

Theorem 4. *Given an $M \times N$ real matrix $\mathbf{A}$ and given an invertible $M \times M$ real matrix $\mathbf{T}$, let $\mathbf{B} = \mathbf{T}\mathbf{A}$, then we have $R_c(\mathbf{B}) = R_c(\mathbf{A})$.*

Proof. Let us consider $\mathbf{x}^T\mathbf{B}^T\mathbf{B}\mathbf{x}$, which is equivalent to $\mathbf{x}^T(\mathbf{A}^T\mathbf{T}^T\mathbf{T}\mathbf{A})\mathbf{x}$ according to the assumption. Suppose that $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum eigenvalue and the maximum eigenvalue of $\mathbf{T}^T\mathbf{T}$, respectively, since $\mathbf{T}^T\mathbf{T}$ is a real symmetric matrix, then by Rayleigh–Ritz theorem [43], we have $\forall \mathbf{x} \in \mathbb{R}^n$, and the inequalities $\lambda_{\min} \cdot \mathbf{x}^T\mathbf{x} \leq \mathbf{x}^T\mathbf{T}^T\mathbf{T}\mathbf{x} \leq \lambda_{\max} \cdot \mathbf{x}^T\mathbf{x}$ hold. Letting $\mathbf{y} = \mathbf{A}\mathbf{x}$, we have $\lambda_{\min} \cdot \mathbf{y}^T\mathbf{y} \leq \mathbf{y}^T\mathbf{T}^T\mathbf{T}\mathbf{y} \leq \lambda_{\max} \cdot \mathbf{y}^T\mathbf{y}$, i.e.,

$$\lambda_{\min} \cdot \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \mathbf{x}^T\mathbf{A}^T\mathbf{T}^T\mathbf{T}\mathbf{A}\mathbf{x} \leq \lambda_{\max} \cdot \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \tag{2}$$

Then, using the left-hand side inequality $\lambda_{\min} \cdot \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \mathbf{x}^T\mathbf{B}^T\mathbf{B}\mathbf{x}$ in Equation (2) and the definition of cover order in Definition 1, we have:

$$\begin{aligned} \left\{\mathbf{x} \in \mathbb{R}_+^N : \mathbf{x}^T\mathbf{B}^T\mathbf{B}\mathbf{x} \leq 1\right\} \quad &\subseteq \quad \left\{\mathbf{x} \in \mathbb{R}_+^N : \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \frac{1}{\lambda_{min}}\right\} \\ &\subseteq \quad \left\{\mathbf{x} \in \mathbb{R}_+^N : x_{k_i} \leq \frac{c_{k_i}}{\lambda_{min}}, i = 1, \cdots, R_c(\mathbf{A})\right\} \end{aligned} \tag{3}$$

where $k_i \in \{1, 2, \cdots, N\}$ and c_{k_i} are positive real numbers. Then, by the definition of cover order in Definition 1, we know that at least $R_c(\mathbf{A})$ variables in $\mathbf{x}$ associated with the column vectors in $\mathbf{B}\mathbf{x}$ are covered. Thus, we have $R_c(\mathbf{A}) \leq R_c(\mathbf{B})$.

According to the right-hand side inequality in Equation (2), we have:

$$\begin{aligned} \left\{\mathbf{x} \in \mathbb{R}_+^N : \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \frac{1}{\lambda_{max}}\right\} \quad &\subseteq \quad \left\{\mathbf{x} \in \mathbb{R}_+^N : \mathbf{x}^T\mathbf{B}^T\mathbf{B}\mathbf{x} \leq 1\right\} \\ &\subseteq \quad \left\{\mathbf{x} : 0 \leq x_{k_i} \leq c_{k_i}, i = 1, \cdots, R_c(\mathbf{B})\right\} \end{aligned} \tag{4}$$

Similarly, at least $R_c(\mathbf{B})$ variables in $\mathbf{x}$ associated with the column vectors in $\mathbf{A}\mathbf{x}$ are covered. Thus, we have $R_c(\mathbf{A}) \geq R_c(\mathbf{B})$.

Hence, we can conclude that, if $\det(\mathbf{T}) \neq 0$ and $\mathbf{B} = \mathbf{T}\mathbf{A}$, then $R_c(\mathbf{A}) = R_c(\mathbf{B})$. □

Theorem 4 carries important implications. It states that the cover order of a matrix is invariant under any row transformation. From Lemma 3, we know that if we are able to find non-negative vectors in $\mathbb{S}_{\mathbf{A}}$, then the cover order of $\mathbf{A}$ is equal to the largest number of the positive entries of these vectors. Thus, Theorem 4 together with Lemma 3 implicitly suggests that we can perform a series of linear elementary row transformations and column permutations to determine the cover order of the matrix. This indeed leads us to the development of a straightforward procedure transforming $\mathbf{A}$ into an *échelon form* for the evaluation of its cover order.

4.2. Procedure of the Échelon Transformation

An échelon form of a rectangular matrix [1] has the following structures:

Definition 3 (échelon form)**.** *A rectangular matrix is in échelon form (or row échelon form) if it has the following three properties:*

(a) *All nonzero rows are above any rows of all zeros.*
(b) *Each leading entry of a row is in a column to the right of the leading entry of the row above it.*
(c) *All entries in a column below a leading entry are zeros.*

Our procedure of échelon transformation can now be laid out as follows:

1. *The Échelon Form of* $\mathbf{A}$. Given an $M \times N$ real matrix $\mathbf{A}$, we can find matrices $\mathbf{E_0}$ and $\mathbf{P_0}$ such that [1]:

$$\mathbf{E}_0\mathbf{A}\mathbf{P}_0 = \begin{pmatrix} \mathbf{I} & \mathbf{B}_0 \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{5}$$

where $\mathbf{I} \in \mathbb{R}^{R_r \times R_r}$ and $\mathbf{B_0} \in \mathbb{R}^{R_r \times (N-R_r)}$ with R_r being the rank of the matrix $\mathbf{A}$. Here, $\mathbf{E_0}$ and $\mathbf{P_0}$ are, respectively, the *elementary transformation* and the *permutation* matrices, either of which may be made up of a product of simpler elementary and permutation matrices. The right side of Equation (5) conforms with the description of *échelon form*; thus, Equation (5) is an *échelon transformation* of $\mathbf{A}$. Note that the échelon transformation of $\mathbf{A}$ is not unique; we can choose different $\mathbf{E}_0$ and $\mathbf{P}_0$, arriving at different values of $\mathbf{B}_0$.

2. *The Cover Order*. Without loss of generality, we can assume that $\mathbf{A}$ has full rank. In particular, from Theorem 1 and Lemma 3, if the initial échelon transformation of $\mathbf{A}$ in Equation (5) results in every entry in some *row* of $\mathbf{B}_0$ being positive, then $R_c(\mathbf{A}) = N$, i.e., $\mathbf{A}$, has full cover. On the other hand, if every entry in some *column* of $\mathbf{B}_0$ is negative, then $R_c(\mathbf{A}) = 0$. However, if the cover order of $\mathbf{A}$ is not immediately obvious from the structure of $\mathbf{B}_0$ resulted from the initial échelon transformation, we need the following steps of structural arrangement to determine the cover order.

(1) *Structure Arrangement*. Search for all non-negative rows in $\mathbf{B}_0$ and select the one which has the greatest number of positive elements. Move this selected row to the first row and assume that it contains N_1 positive entries. By performing the row and column permutation, we can always ensure the identity matrix structure ahead and let the following statements hold:

$$b_{11}, b_{12}, \cdots, b_{1N_1} > 0, b_{1(N_1+1)}, \cdots, b_{1(N-M)} = 0 \tag{6}$$

where $b_{1i}, i = 1, \cdots, N-M$ are the elements in the first row of the new structure of $\mathbf{B}_0$. Ignoring the above first N_1 columns in new $\mathbf{B}_0$, we find all non-negative rows in the remaining part of it and choose the row with the largest number of positive elements. Moving this row to the second row and assuming that it contains N_2 positive entries in the remaining $N - M - N_1$ columns, we have:

$$b_{2(N_1+1)}, b_{2(N_1+2)}, \cdots, b_{2(N_1+N_2)} > 0, b_{2(N_1+N_2+1)}, \cdots, b_{2(N-M)} = 0 \tag{7}$$

where b_{2i}, $i = N_1 + 1, \cdots, N - M$, are the elements in the second row of the new form of $\mathbf{B}_0$ after the above steps. By arranging the following rows similarly, after s times, we obtain:

$$\begin{aligned} b_{11}, b_{12}, \cdots, b_{1N_1} > 0, b_{1(N_1+1)}, \cdots, b_{1(N-M)} = 0 \\ b_{2(N_1+1)}, b_{2(N_1+2)}, \cdots, b_{2(N_1+N_2)} > 0, b_{2(N_1+N_2+1)}, \cdots, b_{2(N-M)} = 0 \\ \vdots \\ b_{s(N_1+N_2+\cdots+N_{s-1}+1)}, \cdots, b_{s(N_1+\cdots+N_s)} > 0, b_{s(N_1+N_2+\cdots+N_s+1)}, \cdots, b_{s(N-M)} = 0 \end{aligned} \tag{8}$$

where b_{ij} in Equation (8), $i = 1, 2, \cdots, s$, $j = 1, 2, \cdots, N - M$, and $s \le M$, are the elements in the first s rows of the structure of $\mathbf{B_0}$ after s times transformation. The procedure ends when one of the following two cases happens:

(a) $\sum_{i=1}^{s} N_i = N - M$, in which case $\mathbf{A}$ has full cover.
(b) There is no non-negative row vector in the row space of the $\mathbf{B}_0$ after s times of transformation.

Let:

$$\bar{\mathbf{B}} = \begin{pmatrix} b_{s+1,N_1+\cdots+N_s+1} & \cdots & b_{s+1,N-M} \\ \vdots & \ddots & \vdots \\ b_{M,N_1+\cdots+N_s+1} & \cdots & b_{M,N-M} \end{pmatrix} \tag{9}$$

(2) *Cover Order*. At the end of the above structural arrangement, we arrive at the conclusion that the cover order of $\mathbf{A}$ is $R_c(\mathbf{A}) = \sum_{i=1}^{s} N_i + s$ and $s \le M$.
The next theorem states the property of the final échelon form of the matrix from which the cover order of $\mathbf{A}$ can be deduced.

Theorem 5. *For any $M \times N$ real matrix $\mathbf{A}$, there exists an elementary matrix $\mathbf{E}$ and a permutation matrix $\mathbf{P}$ such that $\mathbf{EAP} = (\mathbf{I}, \mathbf{B})$, where $\mathbf{I} \in \mathbb{R}^{R_r \times R_r}$, $\mathbf{B} \in \mathbb{R}^{R_r \times (N-R_r)}$, and R_r is the rank of the matrix $\mathbf{A}$. Then $\mathbf{B}$ either:*

1. *Contains at least one non-negative row;*
2. *Contains at least one negative column vector, or there exists one nonpositive column vector, but the same row position where the zero lies will be negative in some other columns of* **B**.

Proof. We shall prove the result by induction. Without loss of generality, we can assume that the matrix **A** has full rank. Indeed, the proof is based on the following steps: (i) Iteration $N-M=1$, i.e., $\mathbf{EAP}=(\mathbf{I},\mathbf{b})$, where $\mathbf{b}\in\mathbb{R}^M$. If **b** contains at least one positive element, then **A** has full cover. If **b** is a negative vector, then **A** has zero cover. If **b** is a nonpositive vector, then the cover order of **A** equals the number of zero terms. (ii) The result holds true for iteration $N-M=K+1$ given that it holds true for iteration $N-M=K$. We prove the desired result in the following:

Suppose that for $N-M=K$, the above conclusion holds, i.e., if $R_c(\mathbf{A})>0$, then **A** can be transformed into $(\mathbf{I},\mathbf{B})$. Let b_{ij} be the ij-th element in **B**, for $i=1,2,\cdots,M$, $j=1,\cdots,K$. We have Equation (8) hold and $R_c(\mathbf{A})=\sum_{i=1}^{s}N_i+s$. In addition, let $\bar{\mathbf{B}}$ be the form in Equation (9) with $K=N-M$, and $\bar{\mathbf{B}}$ either contains at least one negative column or has one nonpositive column, but the position where the zero lies will be negative in some other column of $\bar{\mathbf{B}}$. In the following, we will prove that if the above conclusion holds for $N-M=K$, then this conclusion will also hold when $N-M=K+1$.

When $N-M=K+1$, we can assume that $\mathbf{A}=(\mathbf{I},\mathbf{b}_1,\cdots,\mathbf{b}_K,\mathbf{b}_{K+1})$, where $\mathbf{I}\in\mathbb{R}^{M\times M}$ and $\mathbf{b}_i\in\mathbb{R}^M$, for $i=1,\cdots,K+1$. Without considering the $\mathbf{b}_{K+1}$, we denote the remaining part in **A** as $\bar{\mathbf{A}}$, which equals $(\mathbf{I},\mathbf{b}_1,\cdots,\mathbf{b}_K)$. According to the assumption for $N-M=K$, $\bar{\mathbf{A}}$ can be transformed into échelon form and we let $R_c(\bar{\mathbf{A}})=\sum_{i=1}^{s}N_i+s$. By considering the corresponding $\mathbf{b}_{K+1}$ with the échelon form of $\bar{\mathbf{A}}$ (apply the same row permutation to $\mathbf{b}_{K+1}$ as $\bar{\mathbf{A}}$ permutes in the échelon transformation and we still use $\mathbf{b}_{K+1}$ to denote it after the permutation), we can notice that if $b_{1,K+1}>0$, $R_c(\mathbf{A})=\sum_{i=1}^{s}N_i+s+1$. If $b_{1,K+1}=0$, we can perform the following process from the second row. Therefore, in the following, we will consider the case when $b_{1,K+1}<0$. According to Theorem 4 and Lemma 3, the following steps can be taken to make the first row nonpositive and move it to the last row without affecting the cover order of **A**.

Step 1 Let $m=\max\left\{-\frac{b_{j1}}{b_{11}},\cdots,-\frac{b_{j,K+1}}{b_{1,K+1}}\right\}$, where $b_{j1}b_{11}<0,\cdots,b_{j,K+1}b_{1,K+1}<0$. Multiply the first row of **A** with m, and add the product to i-th row, for $i=2,3,\cdots,M$. We will have $\mathbf{A}_{(1)}$.

Step 2 Use (-1) times the first row of $\mathbf{A}_{(1)}$ to obtain $\mathbf{A}_{(2)}$.

Step 3 Let $t_2=\frac{b_{2,K+1}}{b_{1,K+1}}+m, t_3=\frac{b_{3,K+1}}{b_{1,K+1}}+m,\cdots,t_M=\frac{b_{M,K+1}}{b_{1,K+1}}+m$ and let $\mathbf{a}_j^T$ be the j-th row of $\mathbf{A}_{(2)}$. Then by adding $\mathbf{a}_1^T t_j$ to the j-the row in $\mathbf{A}_{(2)}$, where $j=2,3,\cdots,M$. We will have $\mathbf{A}_{(3)}$.

Step 4 Multiplying the first row of $\mathbf{A}_{(3)}$ with $-\frac{1}{b_{1,K+1}}$ and exchanging the position of the first column with the last column, we will obtain $\mathbf{A}_{(4)}$.

Step 5 Permuting the rows and columns so that the first row in the right-hand side of $\mathbf{A}_{(4)}$ is moved to the last row, as well as securing the left-hand side identity matrix structure. After this, we will have $\mathbf{A}_{(5)}$.

Step 6 Without considering the last column of $\mathbf{A}_{(5)}$, rearranging the rows and columns of the first $(M+K)$ columns of it, we will obtain a new *échelon form* matrix $\bar{\mathbf{A}}_{(5)}$ and $R_c(\bar{\mathbf{A}}_{(5)})=\sum_{i=1}^{s^{(2)}}N_i^{(2)}+s^{(2)}$. By considering the corresponding $\bar{\mathbf{b}}_{K+1}$ with the *échelon form* of $\bar{\mathbf{A}}_{(5)}$, we can notice that if $\bar{b}_{1,K+1}>0$, then $R_c(\mathbf{A})=\sum_{i=1}^{s^{(2)}}N_i^{(2)}+s^{(2)}+1$. If $\bar{b}_{1,K+1}<0$, we can repeat the above steps.

Finally, after either t times transformation, there exists one $b_{1,K+1}>0$, such that the first row of the new matrix is non-negative and $R_c(\mathbf{A})=\sum_{i=1}^{s^{(t)}}N_i^{(t)}+s^{(t)}+1$, or $R_c(\mathbf{A})=0$ and there exists at least one column ($(K+1)$-th column of **A**) which is negative. □

An important indication given by Theorem 5 is that when the first scenario occurs, then the cover order of $\mathbf{A}$ can be determined by the steps as shown in the above échelon transformation; otherwise, $R_c(\mathbf{A}) = 0$.

4.3. Some Properties of the Échelon Form

We observe from the results of échelon transformation that the final échelon form of a matrix is not unique, and under different circumstances, different forms may be required. It is thus interesting to investigate the specific échelon form for special cases, especially for a low-rank matrix $\mathbf{A}$.

1. *Let $\mathbf{A} \in \mathbb{R}^{M\times N}$ and $R_r(\mathbf{A}) = 2$. If $\mathbf{A}$ has full cover, then $\mathbf{A}$ can be transformed into:*

$$\mathbf{A} \rightarrow \begin{pmatrix} \mathbf{I}_2 & \mathbf{B}_{2\times(N-2)} \\ \mathbf{0}_{(M-2)\times 2} & \mathbf{0}_{(M-2)\times(N-2)} \end{pmatrix} \tag{10}$$

where $\mathbf{B}$ is a non-negative matrix.

Proof. Without loss of generality, suppose that $M = R_r(\mathbf{A}) = 2$. If $\mathbf{A}$ has full cover, then by Theorem 5, $\mathbf{A}$ can be transformed into:

$$\begin{pmatrix} 1 & 0 & b_{11} & b_{12} & \cdots & b_{1,(N-2)} \\ 0 & 1 & b_{21} & b_{22} & \cdots & b_{2,(N-2)} \end{pmatrix}$$

where $b_{11}, b_{12}, \cdots, b_{1,(N-2)} > 0$. If $b_{2i} < 0$, where $i \in \{1, \cdots, N-2\}$, let:

$$t = \max_{i\in\{1,\cdots,N-2\}} \left\{ -\frac{b_{2i}}{b_{1i}}, b_{2i} < 0 \right\} = -\frac{b_{2j}}{b_{1j}}$$

Using t times the first row of $\mathbf{A}$ and adding the product to the second row of it, in the next step, multiply the first row of the above matrix with $\frac{1}{b_{1j}}$ and exchange the first column with the j-th column so that the identity matrix structure in the left-hand side part can be guaranteed. After these steps, we obtain a matrix $\mathbf{B}$ which has two non-negative row vectors. □

2. For a rank-2 matrix $\mathbf{A}$, we also have the following property:

Theorem 6. *Let $\mathbf{A} \in \mathbb{R}^{M\times N}$ and $R_r(\mathbf{A}) = 2$. Then $\mathbf{A}$ has zero cover if and only if it can be transformed into the form:*

$$\mathbf{EAP} = \begin{pmatrix} \mathbf{I}_2 & \mathbf{B}^+ & \mathbf{B}^- \\ \mathbf{0}_{(N-2)\times 2} & \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{11}$$

where all the elements in $\mathbf{B}^+$ are non-negative, while the elements in $\mathbf{B}^-$ are all nonpositive. Specifically, $\mathbf{B}^-$ contains at least one column which is a negative vector, or two non-negative vectors with their negative terms lie in different rows.

Proof. Without loss of generality, we can assume that $\mathbf{A}$ has full rank.

Sufficiency: Given a zero-cover matrix $\mathbf{A}$, then by Theorem 5, it can be transformed into $(\mathbf{I}, \mathbf{B})$, with b_{ij} being the ij-th element in $\mathbf{B}$, where $i \in \{1, 2\}$, $j \in \{1, \cdots, N-2\}$. Let $t = \max_{i\in\{1,\cdots,N-2\}}\{-\frac{b_{2i}}{b_{1i}},\ b_{1i}b_{2i} < 0\} = -\frac{b_{2j}}{b_{1j}}$. We multiply the first row of the above transformed matrix with t, and add the product to the second row. Then, we multiply the first row of the resulting matrix with $\frac{1}{b_{1j}}$ and exchange the first column with the j-th column to ensure that the left-hand side part remains an identity matrix. This results in a matrix of the form:

$$\begin{pmatrix} 1 & 0 & \frac{b_{11}}{b_{1j}} & \cdots & \frac{1}{b_{1j}} & \cdots & \frac{b_{1,N-2}}{b_{1j}} \\ 0 & 1 & b_{21} - b_{11}\frac{b_{2j}}{b_{1j}} & \cdots & -\frac{b_{2j}}{b_{1j}} & \cdots & b_{2,N-2} - b_{1,N-2}\frac{b_{2j}}{b_{1j}} \end{pmatrix}$$

If $b_{1i} > 0$ and $i \in \{1, \cdots, N-2\}$, then $b_{2i} - b_{1i}\frac{b_{2j}}{b_{1j}} = b_{1i}\left(\frac{b_{2i}}{b_{1i}} - \frac{b_{2j}}{b_{1j}}\right) > 0$. If $b_{1i} < 0$, then for those $b_{2i} < 0$, we have $b_{2i} - b_{1i}\frac{b_{2j}}{b_{1j}} < 0$, and if $b_{2i} > 0$, the sign of $\left(b_{2i} - b_{1i}\frac{b_{2j}}{b_{1j}}\right)$ is uncertain. After the above steps, and by performing some certain column permutations, the matrix $\mathbf{A}$ can be transformed into the form $\left(\mathbf{I}, \mathbf{B}_{(1)}, \mathbf{B}_{(2)}, \mathbf{B}_{(3)}\right)$, where $\mathbf{B}_{(1)}$ is a non-negative matrix, $\mathbf{B}_{(3)}$ is a nonpositive matrix, and the elements in the first row of $\mathbf{B}_{(2)}$ are all negative, while the elements in the second row of it are all positive. To simplify the discussion, we can write the above matrix as $(\mathbf{I}, \mathbf{B}^{(1)})$, with $b_{ij}^{(1)}$ being the ij-th element in $\mathbf{B}^{(1)}$, where $i \in \{1, 2\}$, $j \in \{1, \cdots, N-2\}$. Then, we let $m = \max_{i \in \{1, \cdots, N-2\}}\{-\frac{b_{1i}^{(1)}}{b_{2i}^{(1)}}, b_{2i}^{(1)} b_{1i}^{(1)} < 0\} = -\frac{b_{1s}^{(1)}}{b_{2s}^{(1)}}$. We multiply the second row of the above matrix by m and add the product to the first row. Then, multiplying the second row of the resulting matrix with $\frac{1}{b_{2s}^{(1)}}$ and exchanging the second column with the s-th column so that the identity matrix structure on the left-hand part can be ensured, we arrive at the following matrix:

$$\begin{pmatrix} 1 & 0 & b_{11}^{(1)} - b_{21}^{(1)}\frac{b_{1s}^{(1)}}{b_{2s}^{(1)}} & \cdots & -\frac{b_{1s}^{(1)}}{b_{2s}^{(1)}} & \cdots & b_{1,N-2}^{(1)} - b_{2,N-2}^{(1)}\frac{b_{1s}^{(1)}}{b_{2s}^{(1)}} \\ 0 & 1 & \frac{b_{21}^{(1)}}{b_{2s}^{(1)}} & \cdots & \frac{1}{b_{2s}^{(1)}} & \cdots & \frac{b_{2,N-1}^{(1)}}{b_{2s}^{(1)}} \end{pmatrix}$$

In the above matrix, we have if $b_{2i}^{(1)} > 0$, where $i \in \{1, \cdots, N-2\}$, then $b_{1i}^{(1)} - b_{2i}^{(1)}\frac{b_{1s}^{(1)}}{b_{2s}^{(1)}} > 0$, and if $b_{2i}^{(1)} < 0$, then $b_{1i}^{(1)} - b_{2i}^{(1)}\frac{b_{1s}^{(1)}}{b_{2s}^{(1)}} < 0$. As a result, if $\mathbf{A}$ has zero cover, then it can be transformed into the right-hand side form in Equation (11).

Necessity: Suppose that $\mathbf{A}$ can be transformed into the form in Equation (11). Then, consider the case when $\mathbf{B}^-$ contains two nonpositive vectors having their negative terms in different rows. Then $\mathbf{Ax} = \mathbf{0}$ can be written as:

$$\begin{pmatrix} \mathbf{I}_{2\times 2} & \mathbf{B}^+ & \bar{\mathbf{B}}^- & \mathbf{b}_1 & \mathbf{b}_2 \\ \mathbf{0}_{(N-2)\times 2} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \times \left(u_1, u_2, \mathbf{x}^+, \mathbf{x}^-, v_1, v_2\right)^T = \mathbf{0}$$

where $\mathbf{b}_1 = (b_1, 0)^T, \mathbf{b}_2 = (0, b_2)^T$, and b_1, b_2 are negative, and $\bar{\mathbf{B}}^-$ is the matrix formed by deleting $\mathbf{b}_1$ and $\mathbf{b}_2$ from $\mathbf{B}^-$. Then we will have:

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} - \mathbf{B}^+\mathbf{x}^+ - \bar{\mathbf{B}}^-\mathbf{x}^-$$

Now, since $\mathbf{x}$ is the solution and must be positive, we can let the elements in $\mathbf{x}^+$ and $\mathbf{x}^-$ take any positive value. If we let v_1 and v_2 be positive and large enough, we can still obtain positive u_1 and u_2. In this case, all elements in $\mathbf{x}$ are positive and satisfy the equation $\mathbf{Ax} = \mathbf{0}$. By Theorem 1, $\mathbf{A}$ has zero cover.

This completes the proof of the theorem. □

5. Cover Order and Linear Programming

In this part, we present a systematic procedure using the concept of hyper-rectangle cover for solving LP problems.

5.1. Linear Programming (LP) Problem

The LP problem [16], in general, can be stated as:

$$\begin{aligned} \min \quad & \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{12}$$

where $\mathbf{A} \in \mathbb{R}^{M \times N}$, with $M < N$, $\mathbf{b} \in \mathbb{R}^M$, and $\mathbf{c}, \mathbf{x} \in \mathbb{R}^N$. We can assume that $\mathbf{A}$ has full rank in general since redundant or inconsistent linear equations can always be detected and removed. The *feasibility set* of the above LP problem is:

$$\mathcal{F}_1 = \left\{ \mathbf{x} \in \mathbb{R}_+^N : \mathbf{A}\mathbf{x} = \mathbf{b} \right\} \subset \mathbb{R}_+^N \tag{13}$$

From the necessary and sufficient condition for the existence of non-negative solution for a nonhomogeneous system of linear equations developed in Theorem 2, we can directly obtain the necessary and sufficient condition that guarantees the nonempty feasibility set of the LP problem. This is stated in the following theorem:

Theorem 7. *The feasibility set $\mathcal{F}_1$ of the LP problem is nonempty iff $R_c(\tilde{\mathbf{A}}) \leq R_c(\mathbf{A})$, where $\tilde{\mathbf{A}} = (\mathbf{A}, -\mathbf{b})$.*

Letting $z = \mathbf{c}^T\mathbf{x}$; then, by adding the objective function into the constraints, the above LP problem can be restated as:

$$\begin{aligned} \min \quad & z \\ \text{subject to} \quad & \begin{pmatrix} \mathbf{A} \\ \mathbf{c}^T \end{pmatrix} \mathbf{x} = \begin{pmatrix} \mathbf{b} \\ z \end{pmatrix} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{14}$$

We denote $\begin{pmatrix} \mathbf{A} \\ \mathbf{c}^T \end{pmatrix}$ as $\mathbf{A}_c$, $\begin{pmatrix} \mathbf{A} & -\mathbf{b} \\ \mathbf{c}^T & -z \end{pmatrix}$ as $\mathbf{A}(z)$. By applying the échelon transformation to $\mathbf{A}(z)$ without changing the position of the last row and the last column, we have:

$$\mathbf{A}(z) \rightarrow \begin{pmatrix} \mathbf{I}_{(M+1)\times(M+1)} & \mathbf{B}_{(M+1)\times(N-M-1)} & \mathbf{f}z + \mathbf{g} \end{pmatrix} \tag{15}$$

where $\mathbf{f}$ and $\mathbf{g}$ are $(M+1) \times 1$ column vectors. To simplify the analysis, in the following, we separate $\mathbf{A}(z)$ into two parts and let:

$$\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{I}_{(M+1)\times(M+1)} & \mathbf{B}_{(M+1)\times(N-M-1)} \end{pmatrix} \text{ and } \tilde{\mathbf{b}} = \mathbf{f}z + \mathbf{g} \tag{16}$$

We have the following observations:

Property 1. *From Theorem 2, in order to have a nonempty feasibility set for this LP problem, adding $\tilde{\mathbf{b}}$ to the right-hand side of $\tilde{\mathbf{A}}$ should not increase the cover order of $\tilde{\mathbf{A}}$. In other words, the cover order of $\mathbf{A}(z)$ should be less than or equal to the cover order of $\mathbf{A}_c$.*

Property 2. *In a minimization problem, if the uncovered variable has a negative coefficient in the objective function and has negative or zero coefficients in all constraints in the échelon form, then the objective function is unbounded over the feasible region.*

5.2. Three Possibilities of the Solution

Based on Property 1, we now analyze the possibilities of the solutions and the optimal value of the objective function of the LP problem under the three conditions: (1) $\tilde{\mathbf{A}}$ has full cover; (2) $0 < R_c(\tilde{\mathbf{A}}) < N$; (3) $\tilde{\mathbf{A}}$ has zero cover, resulting in the following theorem:

Theorem 8. *For the LP problem given by Equation (14): If $\tilde{\mathbf{A}}$ has full cover and the matrix $\mathbf{B}$ in $\tilde{\mathbf{A}}$ is a non-negative matrix, then the LP problem has optimal solution if and only if $\tilde{b}_i = f_i z + g_i \le 0$, $i = 1,\ 2,\ \cdots,\ R_r$. By solving these inequalities, we will have the range of z, which is:*

$$\max\left\{-\frac{g_i}{f_i^{(+)}}, i \in \{1,\cdots,R_r\}\right\} \le z \le \min\left\{-\frac{g_i}{f_i^{(-)}}, i \in \{1,\cdots,R_r\}\right\}, \tag{17}$$

where $f_i^{(+)}$ and $f_i^{(-)}$ are the positive and negative terms in the first R_r elements of $\tilde{\mathbf{b}}$, respectively.

Proof. The proof of the above theorem follows directly from Property 1. □

It should be noted that if the constraint of z in Theorem 8 is contradictory, i.e., if

$$\min\left\{-\frac{g_i}{f_i^{(-)}}, i \in \{1,\cdots,R_r\}\right\} < \max\left\{-\frac{g_i}{f_i^{(+)}}, i \in \{1,\cdots,R_r\}\right\}, \tag{18}$$

then the feasibility set of this linear program is empty: i.e., $\mathcal{F}_1 = \emptyset$. In other words, we are not able to find any feasible solution to this LP problem in this case. In addition, if there is no lower bound of z, i.e., $\max\left\{-\frac{g_i}{f_i^{(+)}}, i \in \{1,\cdots,R_r\}\right\}$ in Equation (17) can be negative infinity, then the objective function in this minimization problem is unbounded.

By the same argument, obtaining the maximum value of z can also be achieved by solving the above inequalities. The maximum value will then be:

$$\max z = \min\{-\frac{g_i}{f_i^{(-)}}, i \in \{1,\cdots,R_r\}\}. \tag{19}$$

If $\tilde{\mathbf{A}}$ has full cover, but the matrix $\mathbf{B}$ is not a non-negative matrix, then let $\mathcal{I} \subseteq \{1,\cdots,R_r\}$ be the index set of the non-negative rows in $\tilde{\mathbf{A}}$. According to the assumption in échelon transformation, the first non-negative row vector in $\tilde{\mathbf{A}}$ contains the largest number of positive terms and the number is N_1. Then the optimal value of the LP problem can be obtained by performing the following steps.

Cover Method (Minimization Form)

Step 1. Solving $f_i z + g_i \le 0$, for $i \in \mathcal{I}$ and a candidate minimal value of z is:

$$z_0 = \max\left\{-\frac{g_i}{f_i^{(-)}}, i \in \mathcal{I}\right\} = -\frac{g_s}{f_s^{(-)}}$$

Step 2. If z_0 satisfies:

$$\max\left\{-\frac{g_k}{f_k^{(-)}}, k \in \{1,\cdots,R_r\}\backslash\mathcal{I}\right\} \le z_0 \le \min\left\{-\frac{g_k}{f_k^{(+)}}, k \in \{1,\cdots,R_r\}\backslash\mathcal{I}\right\}$$

then the process ends and the optimal value is obtained, which is

$$z_{min} = z_0 = \max\left\{-\frac{g_i}{f_i^{(-)}}, i \in \mathcal{I}\right\}$$

Otherwise, there exist some $k \in \{1,\cdots,R_r\}\backslash\mathcal{I}$ such that $f_k z_0 + g_k > 0$, i.e., we have $R_c(\mathbf{A}(z)) > R_c(\mathbf{A}_c)$, then the process continues.

Step 3. Choose column j_k to pivot in (i.e., introduce into the basis variable) by:

$$-\frac{b_{1,j_k}}{b_{k,j_k}} = \min\left\{-\frac{b_{1j}}{b_{kj}}, b_{kj} < 0, 1 \le j \le N_1\right\}$$

Step 4. Choose row $\bar{k}$ to pivot in (i.e., drop from the basis variable) by:

$$\frac{f_{\bar{k}}z_0 + g_{\bar{k}}}{b_{\bar{k},j_k}} = \min\left\{\frac{f_k z_0 + g_k}{b_{kj_k}}, b_{kj} < 0, f_k z_0 + g_k > 0\right\}$$

Step 5. Replace the $\bar{k}$-th column with the $(M + j_k)$-th column and re-establish the échelon form.

Step 6. If the matrix **B** is a non-negative matrix in the new échelon form, then the process ends and the optimal value is obtained, which is

$$z_0 = \max\left\{-\frac{g_i^{new}}{f_i^{(-)new}}, i \in \{1, 2, \cdots, R_r\}\right\}$$

Otherwise, the process continues.

Step 7. Return to step 1.

The whole pivot process each time is performed by using $-\frac{b_{i,j_k}}{b_{\bar{k},j_k}}$ times the $\bar{k}$-th row in $\tilde{\mathbf{A}}$, and adding the product into i-th row, for $i = 1, 2, \cdots, R_r$. Then we divide the $\bar{k}$-th row with $b_{\bar{k},j_k}$, and the $(M + j_k)$-th column becomes $\mathbf{e}_{\bar{k}}$. Next, we exchange the position of the $(M + j_k)$-th column and the $\bar{k}$-th column. After this process, $f_{\bar{k}}z + g_{\bar{k}}$ is negative and the structure of the identity matrix ahead is reserved. The above computational procedures of the cover method in solving the LP problem can be summarized in the flow chart of Figure 3.

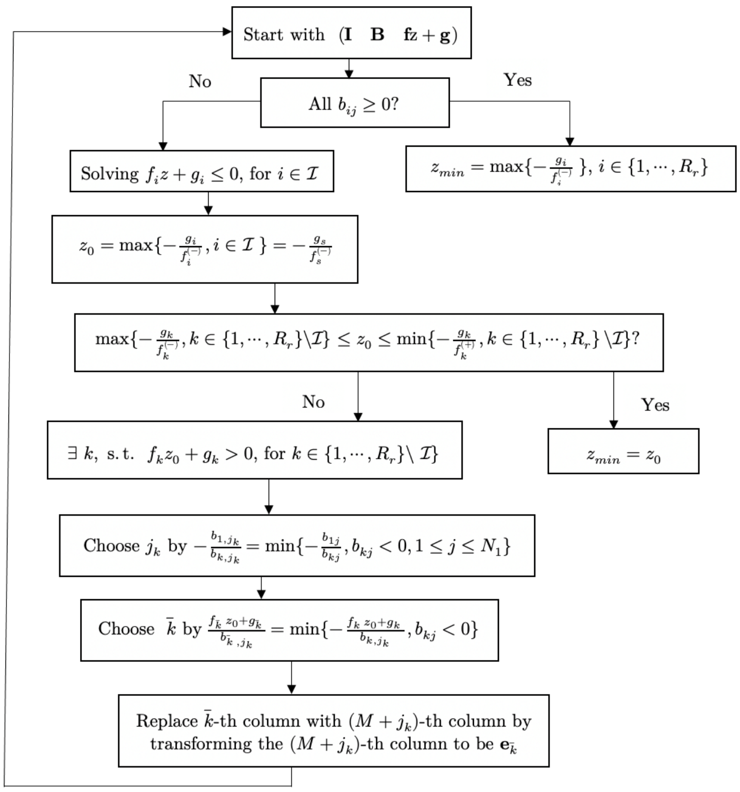

Figure 3. Flow diagram of the cover method.

This simple step-by-step method provides an attractive alternative approach to the LP problem.

The following example provides a clear illustration of the cover method procedure. Here, $\tilde{\mathbf{A}}$ has full cover but the matrix **B** in $\tilde{\mathbf{A}}$ is not a non-negative matrix.

Example 3.

$$\begin{aligned} \min \quad & -x_1 - x_2 \\ & 2x_1 + x_2 + x_3 = 12 \\ & x_1 + 2x_2 + x_4 = 9 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

Letting $z = \mathbf{c}^T\mathbf{x}$ and adding the objective function into the constraints, we will have the following augmented matrix:

$$\tilde{\mathbf{A}}(z) = \begin{pmatrix} 2 & 1 & 1 & 0 & -12 \\ 1 & 2 & 0 & 1 & -9 \\ -1 & -1 & 0 & 0 & -z \end{pmatrix}$$

Applying the échelon transformation to $\tilde{\mathbf{A}}(z)$ without changing the position of the last column, we have

$$\tilde{\mathbf{A}}(z) \to \begin{pmatrix} 1 & 0 & 0 & 1 & -9-z \\ 0 & 1 & 0 & -1 & 9+2z \\ 0 & 0 & 1 & 1 & -21-3z \end{pmatrix}$$

Since we exchange the position of the first two rows during this transformation, the corresponding positions of variables are also exchanged. According to Theorem 8, in order to have the feasible solutions for this LP problem, the following two conditions should be satisfied at the same time: $-9 - z \leq 0$ and $-21 - 3z \leq 0$. By solving these two inequalities, we will have a candidate optimal value of z, which is $z_0 = \max\{-9, -7\} = -7$. Since $z_0 \leq \min\left\{-\frac{g_2}{f_2}\right\} = -\frac{9}{2}$, then the optimal value of the objective function is $z^* = -7$, and the corresponding optimal solution is $\mathbf{x}^* = (5, 2, 0, 0)^T$.

Similarly, for the case when $0 < R_c(\tilde{\mathbf{A}}) < N$, we can also apply the above procedures to obtain the optimal value of the objective function and the optimal solution towards the LP problem by changing the definition of the index set $\mathcal{I}$ and the range of k. For this case, we consider $i \in \mathcal{J}$, and $\mathcal{J} \subseteq \{1, \cdots, s\}$ is the index set of the non-negative rows in the first s rows of $\tilde{\mathbf{A}}$, where s is obtained through the échelon transformation, and $k \in \{1, 2, \cdots, s\} \backslash \mathcal{J}$.

For the zero-cover matrix, the status of the solution for the LP problem is given in the following theorem.

Theorem 9. *For a full rank matrix $\tilde{\mathbf{A}}$, if it has zero cover, then the LP problem is feasible but unbounded.*

Proof. Since adding any column to the right-hand side of a zero-cover matrix will still arrive at a matrix with zero-cover, the feasibility set $\mathcal{F}_1$ is always nonempty in this case. However, according to Theorem 5, a zero-cover matrix can be transformed into a structure which has at least one negative column or has one nonpositive column, but the same row position where the zero lies will be negative in some other column(s) of this structure. Thus, by Property 2, the objective function is unbounded over the feasible domain for the case when $\tilde{\mathbf{A}}$ has zero cover. □

5.3. Feasible Solutions for the LP Problem

With the échelon form and the specific structure of the zero-cover matrix, we are able to obtain a series of feasible solutions for any given LP problem. The detailed process is given in the following:

As we know, $(\tilde{\mathbf{A}}, \tilde{\mathbf{b}})$ is an échelon form matrix, where $\tilde{\mathbf{A}} \in \mathbb{R}^{(M+1)\times N}$ and $\tilde{\mathbf{b}} \in \mathbb{R}^{M+1}$. Then the échelon form can be divided into the following blocks:

$$(\tilde{\mathbf{A}}, \tilde{\mathbf{b}}) = \begin{pmatrix} \mathbf{I}_s & \mathbf{0} & \mathbf{B}^{(1)} & \mathbf{B}^{(3)} & \tilde{\mathbf{b}}^{(1)} \\ \mathbf{0} & \mathbf{I}_{(M+1-s)} & \mathbf{B}^{(2)} & \mathbf{B}^{(4)} & \tilde{\mathbf{b}}^{(2)} \end{pmatrix}, \tag{20}$$

where s is obtained through échelon transformation. Then by Theorem 1, in the system $(\tilde{\mathbf{A}}, \tilde{\mathbf{b}})\mathbf{x} = \mathbf{0}$, $\mathbf{x} \geq \mathbf{0}$, the covered variables x_i are all zeros. As a result, we can ignore those covered column vectors in $(\tilde{\mathbf{A}}, \tilde{\mathbf{b}})$, which correspond to $\begin{pmatrix} \mathbf{I}_s & \mathbf{B}^{(1)} \\ \mathbf{0} & \mathbf{B}^{(2)} \end{pmatrix}$. In Equation (20), since $\mathbf{B}^{(3)}$ is a zero matrix and $\tilde{\mathbf{b}}^{(1)}$ is a zero vector, we only need to consider the remaining part of $(\tilde{\mathbf{A}}, \tilde{\mathbf{b}})$ which is $\left(\mathbf{I}_{(M+1-s)}, \mathbf{B}^{(4)}, \tilde{\mathbf{b}}^{(2)}\right)$. Let us denote this part as $(\bar{\mathbf{A}}, \bar{\mathbf{b}}) = (\mathbf{I}, \bar{\mathbf{B}}, \bar{\mathbf{b}}) = \left(\mathbf{I}_{(M+1-s)}, \mathbf{B}^{(4)}, \tilde{\mathbf{b}}^{(2)}\right)$. The cover order of this matrix is zero. Thus, in order to obtain the feasible solution for the LP problem, we only need to solve the following system of linear equations, where the non-negative vector $\bar{\mathbf{x}}$ is the uncovered part in $\mathbf{x}$:

$$(\mathbf{I}, \bar{\mathbf{B}}, \bar{\mathbf{b}})\bar{\mathbf{x}} = \mathbf{0} \tag{21}$$

For simplicity of discussion, we can assume that $\mathbf{I} \in \mathbb{R}^{m \times m}$, $\bar{\mathbf{B}} \in \mathbb{R}^{m \times (n-m)}$, and $\bar{\mathbf{b}} \in \mathbb{R}^m$. From Theorem 5, we know that the zero-cover matrix can be transformed to the form which contains at least one negative column, or has one nonpositive column, but the row position where the zero lies will be negative in some other column of the matrix. Without loss of generality, we can assume that the negative column appears in the first column of $\bar{\mathbf{B}}$, i.e., $(\bar{b}_{11}, \bar{b}_{21}, \cdots, \bar{b}_{m1})^T$ is a negative column vector. Then the following procedure enables us to obtain a series of feasible solutions to the LP problem.

Suppose that $\bar{\mathbf{x}} = (\bar{x}_1, \cdots, \bar{x}_m, \bar{x}_{m+1}, \bar{x}_{m+2} \cdots, \bar{x}_n, \bar{x}_{n+1})^T$, where $\bar{x}_1, \cdots, \bar{x}_m$ correspond to the column vectors in the $m \times m$ identity matrix, $\bar{x}_{m+1}, \bar{x}_{m+2}, \cdots, \bar{x}_n$ correspond to the column vectors in $\bar{\mathbf{B}}$, and $\bar{x}_{n+1}$ corresponds to $\bar{\mathbf{b}}$ in the multiplication $(\mathbf{I}, \bar{\mathbf{B}}, \bar{\mathbf{b}})\bar{\mathbf{x}}$. Then, by Equation (21), the first m elements in $\bar{\mathbf{x}}$ can be expressed as a linear combination of $\bar{x}_{m+1}, \cdots, \bar{x}_{n+1}$:

$$\begin{aligned} \bar{x}_1 &= -\bar{b}_{11}\bar{x}_{m+1} - \bar{b}_{12}\bar{x}_{m+2} - \cdots - \bar{b}_{1(n-m)}\bar{x}_n - \bar{b}_1\bar{x}_{n+1} \\ &\vdots \\ \bar{x}_m &= -\bar{b}_{m1}\bar{x}_{m+1} - \bar{b}_{m2}\bar{x}_{m+2} - \cdots - \bar{b}_{m(n-m)}\bar{x}_n - \bar{b}_m\bar{x}_{n+1} \end{aligned} \tag{22}$$

In order to obtain a linearly independent feasible solution set, we first let the vector $(\bar{x}_{m+1}, \cdots, \bar{x}_{n+1})^T$ be a set of linearly independent vectors $(L, 1, 0, \cdots, 0)^T$, $(L, 0, 1, \cdots, 0)^T$, $\cdots$, $(L, 0, 0, \cdots, 1)^T$ successively. In addition, in order to satisfy the non-negativity constraints on the variable $\bar{x}_i$, $i = 1, \cdots, n+1$, we let:

$$L = \max\left\{-\frac{\bar{b}_{i2}}{\bar{b}_{11}}, -\frac{\bar{b}_{i3}}{\bar{b}_{21},}, \cdots, -\frac{\bar{b}_{i(n-m)}}{\bar{b}_{m1}}, -\frac{\bar{b}_i}{\bar{b}_{i1}}\right\}, \ i = 1, 2, \cdots, m \tag{23}$$

We can then obtain a set of linear independent basic feasible solutions:

$$\begin{aligned} \boldsymbol{\alpha}_1 &= \left(-\bar{b}_{11}L - \bar{b}_{12}, \cdots, -\bar{b}_{m1}L - \bar{b}_{m2}, L, 1, 0, \cdots, 0\right)^T \\ \boldsymbol{\alpha}_2 &= \left(-\bar{b}_{11}L - \bar{b}_{13}, \cdots, -\bar{b}_{m1}L - \bar{b}_{m3}, L, 0, 1, \cdots, 0\right)^T \\ &\vdots \\ \boldsymbol{\alpha}_{n-m-1} &= \left(-\bar{b}_{11}L - \bar{b}_{1n}, \cdots, -\bar{b}_{m1}L - \bar{b}_{mn}, L, 0, 0, \cdots, 1, 0\right)^T \\ \boldsymbol{\alpha}_{n-m} &= \left(-\bar{b}_{11}L - \bar{b}_1, \cdots, -\bar{b}_{m1}L - \bar{b}_m, L, 0, 0, \cdots, 0, 1\right)^T \end{aligned} \tag{24}$$

Any convex combination of those basic feasible solutions, i.e.,

$$\bar{\mathbf{x}} = k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_{n-m}\boldsymbol{\alpha}_{n-m}$$

where the real numbers k_i satisfy $k_i \geq 0$ and $k_1 + \cdots + k_{n-m} = 1$, is thus a solution of Equation (21). By padding the covered variables into $\bar{\mathbf{x}}$, we obtain a series of feasible solutions to the LP problem.

5.4. The Simplex Method and the Cover Method

In 1947, Dantzig developed an algorithm to solve the LP problem efficiently, called the *simplex method*.

The LP problem is to find the extreme point of this polytope where the objective function is the smallest (or largest) in value if such an extreme point exists. By moving along the edge of the polytope, the simplex method identifies these extreme points with better objective values. The process continues until the optimum objective value is reached, or an unbounded edge is visited. For an LP problem having a nonempty feasible region, the algorithm always terminates because of the finite number of extreme points in the polytope. In practice, the simplex method has shown remarkable efficiency. However, in 1972, Klee and Minty gave an example, the *Klee–Minty cube* [44], showing that the worst-case complexity of the simplex method is exponential time.

While the simplex method regards the objective value z in the canonical tableau of the LP problem as a variable, the *cover method* treats it as a constant. Given a linear program, the cover method first rewrites an LP problem into the form of Equation (14), and then $\mathbf{A}(z)$ is transformed into its échelon form. At this stage, if the matrix $\mathbf{B}$ in this échelon form is a non-negative matrix, then the optimum objective value can be determined directly according to Theorem 8. Thus, the computational complexity of this case is almost entirely determined by the complexity of échelon transformation. In the following, we will review the échelon transformation and analyze its computation complexity.

Consider a full rank matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$ with $M < N$. The complexity of transforming $\mathbf{A}$ into an échelon form is $\mathcal{O}(M^2N)$. In the structure arrangement process of the échelon transformation, the row having the greatest number of nonzero elements is moved to the first row, while the nonzero elements in this row have been moved to the left side of $\mathbf{B}$. Meanwhile, the corresponding column permutation such that the identity matrix structure could be preserved is performed. Thus, the selection of the row having the greatest number of nonzero elements is completed. The next step takes away the columns of $\mathbf{A}$ corresponding to these nonzero elements in the first row and performs an échelon transformation on the remaining part of $\mathbf{A}$. Such an iteration of échelon transformation, each time taking a lower complexity, continues until the desired form is achieved. The complexity of the structural arrangement process is $\mathcal{O}(M^2(N-M))$. Thus, the computation complexity of échelon transformation in solving this type of LP problem by cover method is $\mathcal{O}(M^2N)$.

It is observed, however, that if the matrix $\mathbf{B}$ is not a non-negative matrix, then the cover method for solving the LP problem will involve pivoting steps for which the complexity of the algorithm is no longer polynomial.

6. Cover Length

We first encountered the concept of *cover length* in Definition 1. In this section, we propose a method to determine the cover length of the covered variable x_i associated with the i-th column vector a_i in $\mathbf{Ax}$. In addition, we find a strong relationship between the problem of cover length determination and the *non-negative least square* (NNLS) problem such that we can obtain an analytical result of the NNLS problem by simply determining the cover length of the corresponding variable. We also include a discussion of the various algorithms for solving the NNLS problem and the cover length method developed here.

6.1. Determination of Cover Length

In general, the cover length is obtained by solving the following optimization problem:

Problem 1. *Let* $\mathbf{A} \in \mathbb{R}^{M\times N}$, $\mathbf{x} = \{x_1, x_2, \cdots, x_N\}^T \in \mathbb{R}_+^N$ *and* x_N *be covered in* $\mathbf{Ax}$.

$$\begin{array}{ll} \max & x_N \\ \text{subject to} & \mathbf{x}^T\mathbf{A}^T\mathbf{Ax} \le 1 \end{array} \tag{25}$$

where $x_n \ge 0$ *for* $n = 1, 2, \cdots, N$.

The maximum value of x_N within the constraints is the cover length of the covered variable x_N. To solve the above optimization problem, let us form a Lagrangian function: $L(\mathbf{x}, \boldsymbol{\lambda}) = -x_N - \sum_{n=1}^{N} \lambda_n x_n + \lambda_{N+1}(\mathbf{x}^T\mathbf{A}^T\mathbf{Ax} - 1)/2$, where $\lambda_n > 0$ for $n = 1, 2, \cdots, N+1$. Then, the necessary and sufficient condition for $\mathbf{x}^*$ to be an optimal solution is that the following Karush–Kuhn–Tucker (KKT) condition must be satisfied:

$$\begin{aligned} \nabla L(\mathbf{x}, \boldsymbol{\lambda})|_{\mathbf{x}=\mathbf{x}^*, \boldsymbol{\lambda}=\boldsymbol{\lambda}^*} = -\mathbf{e}_N - \boldsymbol{\lambda}^* + \lambda_{N+1}^*\mathbf{A}^T\mathbf{Ax}^* &= \mathbf{0} \\ x_n^*\lambda_n^* = 0 \qquad \text{for } n = 1, 2, \cdots, &N \\ \lambda_{N+1}^*\big((\mathbf{x}^*)^T\mathbf{A}^T\mathbf{Ax}^* - 1\big) &= 0 \\ (\mathbf{x}^*)^T\mathbf{A}^T\mathbf{Ax}^* &\le 1 \\ \mathbf{x}^* &\ge \mathbf{0} \\ \lambda_{N+1}^* \ge 0,\ \boldsymbol{\lambda}^* &\ge \mathbf{0} \end{aligned} \tag{26}$$

where the non-negative vector $\boldsymbol{\lambda}^* \in \mathbb{R}_+^N$ is associated with the optimal vector $\mathbf{x}^*$ such that $L(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ is a stationary point of $L(\mathbf{x}, \boldsymbol{\lambda})$. On the other hand, we notice that

$$\mathbf{x}^T\mathbf{A}^T\mathbf{Ax} = p_{NN}\Big(x_N + \frac{\bar{\mathbf{p}}_N^T\bar{\mathbf{x}}_N}{p_{NN}}\Big)^2 + \bar{\mathbf{x}}_N^T\Big(\bar{\mathbf{P}}_{NN} - \frac{\bar{\mathbf{p}}_N\bar{\mathbf{p}}_N^T}{p_{NN}}\Big)\bar{\mathbf{x}}_N \tag{27}$$

where $\mathbf{P} = \mathbf{A}^T\mathbf{A}$ is an $N \times N$ positive semidefinite (PSD) matrix, p_{NN} is the NN-th element in $\mathbf{P}$, $\bar{\mathbf{P}}_{NN}$ is the $(N-1) \times (N-1)$ sub-matrix of $\mathbf{P}$ by deleting the N-th row and N-th column from it, $\bar{\mathbf{p}}_N$ is the $(N-1) \times 1$ vector generated by deleting the N-th entry from the N-th row of $\mathbf{P}$, and $\bar{\mathbf{x}}_N$ denotes the $(N-1) \times 1$ vector obtained by deleting the N-th entry from $\mathbf{x}$. Therefore, we can represent the KKT condition alternatively as:

$$\begin{aligned} -\bar{\boldsymbol{\lambda}}_N^* + \lambda_{N+1}^*\Big(\big(x_N^* + \frac{\bar{\mathbf{p}}_N^T\bar{\mathbf{x}}_N^*}{p_{NN}}\big)\bar{\mathbf{p}}_N + \big(\mathbf{P}_{NN} - \frac{\bar{\mathbf{p}}_N\bar{\mathbf{p}}_N^T}{p_{NN}}\big)\bar{\mathbf{x}}_N^*\Big) &= \mathbf{0} \\ -1 - \lambda_N^* + \lambda_{N+1}^*\Big(p_{NN}x_N^* + \bar{\mathbf{p}}_N^T\bar{\mathbf{x}}_N^*\Big) &= 0 \\ x_n^*\lambda_n^* &= 0 \\ \lambda_{N+1}^*\big((\mathbf{x}^*)^T\mathbf{A}^T\mathbf{Ax}^* - 1\big) &= 0 \\ (\mathbf{x}^*)^T\mathbf{A}^T\mathbf{Ax}^* &\le 1 \\ \mathbf{x}^* &\ge \mathbf{0} \\ \lambda_{N+1}^* \ge 0,\ \boldsymbol{\lambda}^* &\ge \mathbf{0} \end{aligned} \tag{28}$$

Here, $\bar{\boldsymbol{\lambda}}_N^*$ and $\bar{\mathbf{x}}_N^*$ denote, respectively, the $(N-1) \times 1$ vectors obtained by deleting the N-th entry from $\boldsymbol{\lambda}^*$ and $\mathbf{x}^*$. Since $x_N^* \neq 0$, we have $\lambda_{N+1}^* \neq 0$, and, thus, $x_N^* = \lambda_{N+1}^*$. Using the KKT condition, the solution to Problem 1 is given in the following theorem:

Theorem 10. *Let* $\mathbf{A}$ *be an* $M \times N$ *real matrix with its rank being* R. *Then,* x_n *is covered in* $\mathbf{Ax}$ *if and only if there exists an invertible principal sub-matrix* $\mathbf{P}_{i_1 i_2 \cdots i_r}$ *of order* r *in* $\mathbf{A}^T\mathbf{A}$ *that includes the nn-th element* $[\mathbf{A}^T\mathbf{A}]_{nn}$, *such that the following two conditions are satisfied simultaneously:*

1. $\mathbf{P}_{i_1 i_2 \cdots i_r | i_j = n}^{-1}\mathbf{e}_j \ge \mathbf{0}$ *and* $\Big[\mathbf{P}_{i_1 i_2 \cdots i_r | i_j = n}^{-1}\mathbf{e}_j\Big]_j > 0$;

2. $\det(\mathbf{P}_{i_1\cdots i_r|i_j=n\to k}) \geq 0$ *for* $k \in \{1,\cdots,N\}\backslash\{i_1,\cdots,i_r\}$*, where* $\mathbf{P}_{i_1\cdots i_r|i_j=n\to k}$ *is the sub-matrix of* $\mathbf{A}^T\mathbf{A}$ *by replacing the old row :* $(p_{n,i_1},\cdots,p_{n,i_r})$ *in* $\mathbf{P}_{i_1\cdots i_r|i_j=n}$ *with the new row* $(p_{k,i_1},\cdots,p_{k,i_r})$*.*

   Then the cover length is given by $c_n(x_n) = \sqrt{\left[\left(\mathbf{P}_{i_1 i_2\cdots i_r|i_j=n}\right)^{-1}\right]_{nn}}$*.*

Proof. The KKT condition of Problem 1 can be simplified as: $\mathbf{Px} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}, \mathbf{b} \geq \mathbf{0}$, $b_N > 0$ and $x_i b_i = 0$ for $i = 1,2, \cdots, N-1$, where $\mathbf{b} \in \mathbb{R}_+^N$. Let $\bar{\mathcal{N}}$ be the set consisting of all the indices of x_i which are all positive in the variable $\mathbf{x}$. Denoting the cardinality of a set as $|\cdot|$, we are able to find an $|\bar{\mathcal{N}}| \times |\bar{\mathcal{N}}|$ sub-matrix $\bar{\mathbf{P}}$ of $\mathbf{P}$, such that $\bar{\mathbf{P}}\bar{\mathbf{x}} = \mathbf{e}_{|\bar{\mathcal{N}}|}$ with all x_i in $\bar{\mathbf{x}}$ being uncovered variables in $\mathbf{x}$ and $\bar{x}_{|\bar{\mathcal{N}}|} = x_N$, i.e., the last entry in $\bar{\mathbf{x}}$ is equivalent to the last one in $\mathbf{x}$. Then there exists a full column rank matrix $\mathbf{T} \in \mathbb{R}^{|\bar{\mathcal{N}}|\times r}$, where $r \leq |\bar{\mathcal{N}}|$, containing $\bar{\mathbf{P}}_{|\bar{\mathcal{N}}|}$ in $\bar{\mathbf{P}}$. Without loss of generality, we can let $\mathbf{T} = \left\{\mathbf{t}_1,\cdots,\mathbf{t}_{r-1},\bar{\mathbf{P}}_{|\bar{\mathcal{N}}|}\right\}$ and we will have $\mathbf{T}\tilde{\mathbf{x}} = \mathbf{e}_{|\bar{\mathcal{N}}|}$, where $\tilde{\mathbf{x}} \in \mathbb{R}_+^r$ and $\tilde{x}_r > 0$. This can be proved in the following: Let $\mathbf{T}$ be the smallest set containing $\bar{\mathbf{P}}_{|\bar{\mathcal{N}}|}$, s.t. $\mathbf{e}_{|\bar{\mathcal{N}}|} \in \text{cone}\,\mathbf{T}$, where cone $\mathbf{T} = \text{cone}\left\{\mathbf{t}_1,\cdots,\mathbf{t}_{r-1},\bar{\mathbf{P}}_{|\bar{\mathcal{N}}|}\right\} = \{\theta_1\mathbf{t}_1 + \cdots + \theta_{r-1}\mathbf{t}_{r-1} + \theta_r\bar{\mathbf{P}}_{|\bar{\mathcal{N}}|} | \theta_i \geq 0$ for $i = 1,\cdots,r\}$. Then $\mathbf{T}$ is linearly independent; otherwise, there are $\mu_j \in \mathbb{R}$ (not all 0), s.t. $\sum_{j=1}^{r-1}\mu_j\mathbf{t}_j + \mu_r\bar{\mathbf{P}}_{|\bar{\mathcal{N}}|} = \mathbf{0}$. And there are $\lambda_j \geq 0$, s.t. $\sum_{j=1}^{r-1}\lambda_j\mathbf{t}_j + \lambda_r\bar{\mathbf{P}}_{|\bar{\mathcal{N}}|} = \mathbf{e}_{|\bar{\mathcal{N}}|}$, where $\lambda_r > 0$. Then, we have $\sum_{j=1}^{r-1}(\alpha\mu_j + \lambda_j)\mathbf{t}_j + (\alpha\mu_r + \lambda_r)\bar{\mathbf{P}}_{|\bar{\mathcal{N}}|} = \mathbf{e}_{|\bar{\mathcal{N}}|}$. If $\mu_r \geq 0$, then let $\alpha = \max_{1\leq j\leq r-1}\{-\frac{\lambda_j}{\mu_j}, \mu_j > 0\} = -\frac{\lambda_i}{\mu_i}$. Thus, for every $1 \leq j \leq r-1$, $\lambda_j + \alpha\mu_j \geq 0$ and $\lambda_i + \alpha\mu_i = 0$. Then, we can have a new $\tilde{\mathbf{x}} \in \mathbb{R}^{r-1}$ with the last element of it, which is $\lambda_r + \alpha\mu_r$, being positive, while the others are $r-2$ elements, which are expressed as $\lambda_j + \alpha\mu_j$, being non-negative. When $\mu_r < 0$, we let $\alpha = \max_{1\leq j\leq r-1}\{-\frac{\lambda_j}{\mu_j}, \mu_j > 0\} = -\frac{\lambda_i}{\mu_i}$. Then, we have a new $\tilde{\mathbf{x}} \in \mathbb{R}^{r-1}$ with the last element of it being positive while the others are non-negative in the same manner as the case when $\mu_r \geq 0$. As a result, we can always find a smaller set $\tilde{\mathbf{T}}$ containing $\bar{\mathbf{P}}_{|\bar{\mathcal{N}}|}$, s.t. $\mathbf{e}_{|\bar{\mathcal{N}}|} \in \text{cone}\,\tilde{\mathbf{T}}$. Thus, $\mathbf{T}$ is linearly independent. According to the constraints of $\mathbf{x}$, $\tilde{\mathbf{x}}$ should be equivalent to $\bar{\mathbf{x}}$ and $\mathbf{T} = \bar{\mathbf{P}}$. As a result, $\bar{\mathbf{P}}$ is invertible. Since we have $\bar{\mathbf{P}}\bar{\mathbf{x}} = \mathbf{e}_{|\bar{\mathcal{N}}|}$, where $\bar{\mathbf{x}} \in \mathbb{R}_+^r$ and $\bar{x}_r > 0$, we will have $\mathbf{P}_{i_1 i_2\cdots i_r|i_j=n}^{-1}\mathbf{e}_j \geq \mathbf{0}$ and the j-th element in it is positive. Until now, the first statement has been proven.

By using the new row $(p_{k,i_1},\cdots,p_{k,i_r})$ to replace the old row $(p_{n,i_1},\cdots,p_{n,i_r})$ in $\mathbf{P}_{i_1 i_2\cdots i_r|i_j=n}$, we will have $\mathbf{P}_{i_1 i_2\cdots i_r|i_j=n\to k}\bar{\mathbf{x}} = b_k\mathbf{e}_j$. To simplify the expression, we denote $\mathbf{P}_{i_1 i_2\cdots i_r|i_j=n\to k}$ as $\bar{\mathbf{P}}_k$. If $\bar{\mathbf{P}}_k$ is invertible, then $x_n = \frac{b_k|\bar{\mathbf{P}}_{(r-1)\times(r-1)}|}{|\bar{\mathbf{P}}_k|}$, where $\bar{\mathbf{P}}_{(r-1)\times(r-1)}$ is the $(r-1)$-th order leading principle sub-matrix of $\mathbf{P}_{i_1 i_2\cdots i_r}$. Since $\mathbf{P}_{i_1 i_2\cdots i_r}$ is a positive definite matrix, it follows that $\left|\bar{\mathbf{P}}_{(r-1)\times(r-1)}\right| > 0$. In addition, $b_k \geq 0$, $x_n > 0$, then we have $\det(\bar{\mathbf{P}}_k) > 0$. When $b_k = 0$, $\det(\bar{\mathbf{P}}_k) = 0$. As a result, $\det(\bar{\mathbf{P}}_k) \geq 0$, for $k = 1,2,\cdots,N$ but $k \neq i_1, i_2, \cdots, i_r$.

When the above conditions are all satisfied, the cover length of x_n can be obtained directly, which is $\sqrt{\left[\left(\mathbf{P}_{i_1 i_2\cdots i_r|i_j=n}\right)^{-1}\right]_{nn}}$. □

From the above result, we can also conclude that if there is no principal sub-matrix that can satisfy all the conditions, then the corresponding variable is uncovered within the constraint.

The following example illustrates how the above method can be used to obtain the cover length of the covered variable:

Example 4. *Determine the cover length of covered variable x_4 given the following 4×4 matrix and its PSD matrix:*

$$\mathbf{A} = \begin{pmatrix} -3 & -2 & -5 & -2 \\ 3 & -5 & 0 & -4 \\ 1 & 3 & 1 & -3 \\ 2 & 2 & 1 & 4 \end{pmatrix}, \quad \mathbf{P} = \mathbf{A}^T\mathbf{A} = \begin{pmatrix} 23 & -2 & 18 & -1 \\ -2 & 42 & 15 & 23 \\ 18 & 15 & 27 & 11 \\ -1 & 23 & 11 & 45 \end{pmatrix}$$

We need to find out the principal sub-matrix that can satisfy all the conditions listed in Theorem 10. We first examine all the 2×2 principal sub-matrices of $\mathbf{P}$ containing a negative element in the right upper side corner, since only such a principle sub-matrix of order-2 could satisfy the condition that the last column of its inverse is a positive vector. Inspection of the above PSD matrix shows that there is only one such principal sub-matrix: $\mathbf{P}_{14} = \begin{pmatrix} 23 & -1 \\ -1 & 45 \end{pmatrix}$. We verify that $\mathbf{P}_{14}$ above is invertible and the last column of its inverse matrix is a positive column vector. Then, we replace the second row in $\mathbf{P}_{14}$ with the other rows, resulting in $\mathbf{P}_{14|4\to 2} = \begin{pmatrix} 23 & -1 \\ -2 & 23 \end{pmatrix}$ and $\mathbf{P}_{14|4\to 3} = \begin{pmatrix} 23 & -1 \\ 18 & 11 \end{pmatrix}$. The determinants of both are verified to be non-negative. From the above discussion, we can see that the invertible 2×2 principal sub-matrix $\mathbf{P}_{14}$ satisfies all the conditions in Theorem 10 and we have: $\mathbf{P}_{14}^{-1} = \begin{pmatrix} \frac{45}{1034} & \frac{1}{1034} \\ \frac{1}{1034} & \frac{23}{1034} \end{pmatrix}$. Thus, we can conclude that the cover length of x_4 is $c_4(x_4) = \sqrt{\frac{23}{1034}}$.

Lemma 4. *For any $\mathbf{A} \in \mathbb{R}^{M \times N}$ and $\mathbf{x} \in \mathbb{R}_+^N$,*

1. *If all the entries of $\mathbf{A}^T\mathbf{A}$ are positive, then the cover length of x_n is $c_n(x_n) = \frac{1}{\sqrt{[\mathbf{A}^T\mathbf{A}]_{nn}}}$.*
2. *If $\mathbf{A}^T\mathbf{A}$ has full rank and all the entries in the n-th column of $(\mathbf{A}^T\mathbf{A})^{-1}$ are positive, then the cover length is $c_n(x_n) = \sqrt{\left[(\mathbf{A}^T\mathbf{A})^{-1}\right]_{nn}}$.*

Proof. To prove the first statement, given an $M \times N$ real matrix $\mathbf{A}$ and $\mathbf{x} \in \mathbb{R}_+^N$, we can rewrite $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x}$ as:

$$\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = \bar{\mathbf{x}}^T\bar{\mathbf{A}}^T\bar{\mathbf{A}}\bar{\mathbf{x}} + \bar{\mathbf{x}}^T\bar{\mathbf{A}}^T\mathbf{a}_n x_n + \mathbf{a}_n^T\bar{\mathbf{A}}\bar{\mathbf{x}}x_n + \mathbf{a}_n^T\mathbf{a}_n x_n^2$$

where $\bar{\mathbf{A}}$ is the $M \times (N-1)$ sub-matrix formed by deleting the n-th column of $\mathbf{A}$, $\bar{\mathbf{x}}$ denotes an $(N-1) \times 1$ vector obtained by deleting n-th entry from $\mathbf{x}$ and $\mathbf{a}_n$ is the n-th column of $\mathbf{A}$.

According to the assumption of Statement 1, i.e., all the entries of $\mathbf{A}^T\mathbf{A}$ are positive, then all terms in the above equations are non-negative. Thus, for any given positive real-valued number $\tau > 0$, $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \le \tau^2$ implies that $\mathbf{a}_n^T\mathbf{a}_n x_n^2 \le \tau^2$, which gives $x_n \le \frac{\tau}{\sqrt{\mathbf{a}_n^T\mathbf{a}_n}}$. Therefore, according to the definition of cover length in Definition 1, the cover length of x_n is given by $c_n(x_n) = \frac{1}{\sqrt{[\mathbf{A}^T\mathbf{A}]_{nn}}}$.

The second statement can be obtained from Theorem 10 directly. □

6.2. Cover Length Problem and NNLS Problem

The NNLS problem is a constrained least squares regression problem in which all the variables can only take non-negative values. Specifically, the NNLS problem can be stated as follows [45]:

Problem 2 (Non-negative Least Squares (NNLS))**.** *Given $\mathbf{B} \in \mathbb{R}^{M \times N}$ and $\mathbf{b} \in \mathbb{R}^M$, find a non-negative vector $\mathbf{u} \in \mathbb{R}_+^N$ such that*

$$\begin{aligned} \min \quad & \| \mathbf{B}\mathbf{u} - \mathbf{b} \|_2^2 \\ \text{subject to} \quad & \mathbf{u} \ge \mathbf{0} \end{aligned} \tag{29}$$

In the following, we show by introducing a new variable that the NNLS problem can be turned into a problem of determining the cover length of the corresponding variable. In so doing, a connection between cover length determination and the NNLS problem is established, providing us with a method to arrive at the closed-form optimal value of the objective function.

First, we let

$$\tau^2 = \| \mathbf{Bu} - \mathbf{b} \|_2^2 \tag{30}$$

When $\tau = 0$, Problem 2, is equivalent to the problem of finding solutions for the nonhomogeneous system of linear equations $\mathbf{Bu} = \mathbf{b}$ with non-negative constraints on $\mathbf{u}$. Let us consider the case when $\tau > 0$: by dividing τ^2 on both sides of Equation (30), we will have

$$\| \mathbf{B}\frac{\mathbf{u}}{\tau} - \mathbf{b}\frac{1}{\tau} \|_2^2 = 1 \tag{31}$$

Introducing a new variable $\mathbf{x} = \left(\frac{\mathbf{u}}{\tau}, \frac{1}{\tau}\right)^T$, the origin problem can be transformed into:

Problem 3.

$$\begin{array}{rl} \max & x_{N+1} \\ \text{subject to} & \| \mathbf{Ax} \|_2^2 = 1 \end{array} \tag{32}$$

where $x_n \geq 0$ *for* $n = 1, 2, \cdots, N$, $x_{N+1} = \frac{1}{\tau} > 0$ *and* $\mathbf{A} = (\mathbf{B}, -\mathbf{b})$.

We observe that Problem 3 is of the same form as Problem 1 and is consistent with Problem 2. Thus, the NNLS and the cover length determination problem are equivalent. By solving the cover length of the corresponding variable x_{N+1}, we obtain the equivalent closed-form optimal value of the objective function in the NNLS problem. If we are not able to find the cover length of this variable, x_{N+1} is unbounded within the constraint and the optimum value of the objective function in the NNLS problem is almost zero.

Example 5. *The cover length determination problem in Example 4 is consistent with the NNLS problem:* $\min_{\mathbf{u}\in\mathbb{R}^3_+} \| \mathbf{Bu} - \mathbf{b} \|_2^2$, *where* $\mathbf{B} = \begin{pmatrix} -3 & -2 & -5 \\ 3 & -5 & 0 \\ 1 & 3 & 1 \\ 2 & 2 & 1 \end{pmatrix}$ *and* $\mathbf{b} = (2, 4, 3, -4)^T$. *Let* $\tau^2 = \| \mathbf{Bu} - \mathbf{b} \|_2^2$ *and* $\mathbf{x} = (x_1, x_2, x_3, x_4)^T = (\frac{\mathbf{u}}{\tau}, \frac{1}{\tau})^T$. *The cover length of* x_4 *is* $c_4(x_4) = \frac{1}{\sqrt{23/1034}} = \frac{1}{\tau}$; *thus, the optimal value of this NNLS problem is* $\tau^2 = \| \mathbf{Bu} - \mathbf{b} \|_2^2 = \left(\frac{1}{c_4(x_4)}\right)^2 = 23/1034$.

The above example demonstrates how to convert the cover length determination of a desired variable into the optimal value of the corresponding NNLS problem and verifies the equivalence of the two problems. For certain types of matrices, using this equivalence, we can even *directly* obtain the analytical optimal value of the NNLS problem. This is demonstrated by the example of the *M-matrix* in the following. Let us first define the Z- and the M-matrices [46]:

Definition 4 (Z-matrix)**.** *An* $N \times N$ *real matrix in which the off-diagonal entries are less than or equal to zero, i.e., a matrix of the form* $\mathbf{A} = (a_{ij})$ *with* $a_{ij} \leq 0 \; \forall \; i \neq j, 1 \leq i, j \leq N$, *is a real Z-matrix.*

Definition 5 (M-matrix)**.** *Let* $\mathbf{A}$ *be an* $N \times N$ *real Z-matrix. Then* $\mathbf{A}$ *is also an M-matrix if it can be expressed in the form* $\mathbf{A} = s\mathbf{I} - \mathbf{T}$, *where* $\mathbf{T} = (t_{ij})$ *with* $t_{ij} \geq 0$, *for all* $i \neq j$, $1 \leq i, j \leq N$, *where s is at least as large as the maximum of the moduli of the eigenvalues of* $\mathbf{T}$*, and* $\mathbf{I}$ *is an identity matrix.*

Theorem 11. *Let* $\mathbf{A} \in \mathbb{R}^{N\times N}$ *be a Z-matrix; then, the following statements are equivalent to* $\mathbf{A}$ *being a nonsingular M-matrix:*

1. *All the principal minors of* $\mathbf{A}$ *are positive. That is, the determinant of each sub-matrix of* $\mathbf{A}$ *obtained by deleting a set, possibly empty, of corresponding rows and columns of* $\mathbf{A}$ *is positive.*
2. $\mathbf{A}$ *is inverse-positive. That is,* $\mathbf{A}^{-1}$ *exists and* $\mathbf{A}^{-1}$ *is a non-negative matrix.*

Then, with the properties of M-matrix and cover length, we have the following result.

Theorem 12. *Let matrix* $\mathbf{B} \in \mathbb{R}^{N\times(N-1)}$ *and vector* $\mathbf{b} \in \mathbb{R}^N$*. Denote* $\mathbf{A}$ *as* $(\mathbf{B}, -\mathbf{b})$*. Supposing that* $\mathbf{A}$ *is a nonsingular M-matrix, then the optimal value of the NNLS problem* $\min\limits_{\mathbf{u}\in\mathbb{R}_+^{N-1}} \parallel \mathbf{Bu} - \mathbf{b} \parallel_2^2$ *is exactly equal to* $\frac{1}{\left[(\mathbf{A}^T\mathbf{A})^{-1}\right]_{NN}}$.

Proof. By assumption, $\mathbf{A}$ is a nonsingular M-matrix, therefore $\mathbf{A}^T\mathbf{A}$ is invertible and all the elements in $(\mathbf{A}^T\mathbf{A})^{-1}$ are positive. Reformulating the NNLS problem to the problem of determining the cover length of x_N by applying Lemma 4, the cover length of the corresponding variable x_N is given by $c_N(x_N) = \sqrt{\left[(\mathbf{A}^T\mathbf{A})^{-1}\right]_{NN}}$. Thus, the optimal value $\tau^2 = (\frac{1}{c_N(x_N)})^2 = \frac{1}{\left[(\mathbf{A}^T\mathbf{A})^{-1}\right]_{NN}}$. □

6.3. Comparison with the Active-Set Method

There are several normally used active-set methods for solving the NNLS problem. A typical example is the algorithm `lsqnonneg` in Matlab, which aims at creating an active set and using it to arrive at an approximate solution to the NNLS problem. `lsqnonneg` starts with an all-zero vector and computes the associated negative gradient vector $\mathbf{w}$. Then it finds the index of the position where the maximum value in $\mathbf{w}$ occurs and moves this index from the inactive set to the active set. By solving the corresponding least squares problem with the current active set, one non-negative solution candidate can be obtained. The active set and inactive set can be updated with the current candidate solution and continue the whole process until all the elements in $\mathbf{w}$ are nonpositive or the inactive set is empty. As Lawson and Hanson showed, this algorithm always converges and terminates in finite steps. However, there is no upper limit on the possible number of iterations that the algorithm might need to reach the point of the optimum solution, and it might be very slow in practice, owing largely to the computation of the pseudo-inverse. With regard to the computational complexity, since the exact running time required for the NNLS solver is unknown, the computational cost cannot be specified exactly. In many standard implementations of NNLS solvers (and particularly those based on active-set methods), the cost is typically $\mathcal{O}(MN^2)$ per iteration [47].

Compared with the active-set method, the cover length determination method is finite, and once we find one principal sub-matrix that can satisfy the conditions in Theorem 10, then the computation stops. Furthermore, we can find an upper limit on the possible number of steps that the algorithm needs and obtain a closed-form optimal value of the NNLS problem.

From the perspective of computation complexity, there is no clear advantage of the cover length method compared with the `lsqnonneg` since it involves the combination and permutation operations. However, while the accuracy of the `lsqnonneg` solution depends on a prescribed tolerance ϵ, the cover length method yields the exact optimal value of the objective function.

We now present some numerical results illustrating the performance of the cover length method and `lsqnonneg` in solving the NNLS problem.

Table 1 shows the average running time (seconds) and average error of the `lsqnonneg` and the cover length method for the matrices and vectors randomly generated by Matlab's `rand` function. The results shown here are averaged over 100 random samples with varying number of columns (from one to three) of **B** in the NNLS problem. The default termination tolerance on the solution of `lsqnonneg` is $10 \times \sum_{ij} |a_{ij}| \times N \times eps$, where $eps = 2.22 \times 10^{-16}$, N is the row number of the matrix **B**, and a_{ij} is the element in $\mathbf{A} = (\mathbf{B}, -\mathbf{b})$. Table 1 also includes the computation complexity (number of maximum operations) of the cover length method in solving the NNLS problem. It is clear from the table that the advantage of the cover length method over `lsqnonneg` lies in the accuracy of the optimal value since cover length yields a closed-form one.

Table 1. Comparison between Matlab's `lsqnonneg` and cover length method in solving the NNLS problem.

Column Number of B		**1**	**2**	**3**
Complexity	Cover length method	6	16	589
Running time	Cover length method	1.20×10^{-4}	2.40×10^{-4}	0.0024
	`lsqnonneg`	2.10×10^{-4}	3.17×10^{-4}	4.18×10^{-4}
Average error	Cover length method	0.0000	0.0000	0.0000
	`lsqnonneg`	1.1102×10^{-16}	2.6645×10^{-15}	2.8422×10^{-14}

7. Conclusions

Linear systems of equations with non-negativity constraints on solutions is an area of study in linear algebra. Such problems arise frequently in many fields of science and engineering. In our consideration of such problems, we discovered the hyper-rectangle cover theory of a matrix, which is presented in this paper. The two main concepts in the hyper-rectangle cover theory, viz., the cover order and the cover length, were first defined, and many of their important properties were introduced. Based on this theory, several novel approaches to analyzing the above typical problems were proposed. The necessary and sufficient conditions under which a unique solution for a system of linear equations with non-negativity constraints exists were identified. We also showed how the specific échelon form of the matrix is constructed, and with this échelon form, the cover order of any given matrix can be determined.

With the help of cover theory, the emptiness of the feasibility set and the various possibilities of the solution for the LP problem were analyzed in detail. In addition, with the property of zero-cover matrix, a series of feasible solutions to the LP problem can be obtained.

Our study on the cover length led us to the development of a method to find the cover length of a covered variable. We also showed the equivalence between cover length determination and the NNLS problem so that the NNLS problem can be solved with the cover length method. This provides us with the analytical optimal value obtainable from the structure of the matrix rather than a numerical result having a finite accuracy. The development of the hyper-rectangle cover theory, thus, not only provides us with an efficient method to solve the system of linear equations with non-negativity constraints, it also suggests to us attractive alternative approaches to the LP and the NNLS problems.

Author Contributions: Conceptualization, X.C., K.M.W. and J.Z.; methodology, X.C., K.M.W., J.C. and J.Z.; software, X.C.; validation, X.C. and K.M.W.; formal analysis, X.C., K.M.W. and J.Z.; investigation, X.C., K.M.W. and J.Z.; resources, X.C. and J.Z.; data curation, X.C.; writing—original draft preparation, X.C.; writing—review and editing, K.M.W., J.C. and J.Z.; visualization, X.C. and K.M.W.; supervision, K.M.W.; project administration, X.C. and K.M.W.; funding acquisition, K.M.W. and J.Z.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: This paper was written with profound gratitude to, and in fond memory of, Jiankang Zhang who, while pioneering much of the original research in hyper-rectangle cover theory, guided the first author (X.C.) with the utmost care in her research, and introduced, with great enthusiasm, some intriguing ideas in the subject to the second author (K.M.W.). His sudden departure from life left those who worked with him with a feeling of loss. He will always be remembered as a great teacher and a great colleague.

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

References

1. Lay, D.C. *Linear Algebra and Its Applications*, 5th ed.; Pearson: New York, NY, USA, 2016.
2. Bebikhov, Y.V.; Semenov, A.; Yakushev, I.; Kugusheva, N.; Pavlova, S.; Glazun, M. The application of mathematical simulation for solution of linear algebraic and ordinary differential equations in electrical engineering. In Proceedings of the IOP Conference Series: Materials Science and Engineering, Wuhan, China, 10–12 October 2019; IOP Publishing: Bristol, UK, 2019; Volume 643, p. 012067.
3. Dianat, S.A.; Saber, E. *Advanced Linear Algebra for Engineers with MATLAB*; CRC Press: Boca Raton, FL, USA, 2017.
4. Golomb, S.W.; Gong, G. *Signal Design for Good Correlation: For Wireless Communication, Cryptography, and Radar*; Cambridge University Press: Cambridge, UK, 2005.
5. Bardsley, J.M.; Knepper, S.; Nagy, J. Structured linear algebra problems in adaptive optics imaging. *Adv. Comput. Math.* **2011**, *35*, 103. [CrossRef]
6. Datta, B.N. Linear and numerical linear algebra in control theory: Some research problems. *Linear Algebra Its Appl.* **1994**, *197*, 755–790. [CrossRef]
7. Joshi, H.; Yavuz, M.; Townley, S.; Jha, B.K. Stability analysis of a non-singular fractional-order COVID-19 model with nonlinear incidence and treatment rate. *Phys. Scr.* **2023**, *98*, 045216. [CrossRef]
8. Joshi, H.; Jha, B.K.; Yavuz, M. Modelling and analysis of fractional-order vaccination model for control of COVID-19 outbreak using real data. *Math. Biosci. Eng.* **2023**, *20*, 213–240. [CrossRef]
9. Joshi, H.; Jha, B.K. 2D memory-based mathematical analysis for the combined impact of calcium influx and efflux on nerve cells. *Comput. Math. Appl.* **2023**, *134*, 33–44. [CrossRef]
10. Anton, H.; Rorres, C. *Elementary Linear Algebra: Applications Version*; John Wiley & Sons: Hoboken, NJ, USA, 2013.
11. Demmel, J.W. Matrix Computations (Gene H. Golub and Charles F. van Loan). *SIAM Rev.* **1986**, *28*, 252–255. [CrossRef]
12. Horn, R.A.; Johnson, C.R. *Matrix Analysis*; Cambridge University Press: Cambridge, UK, 2012.
13. Roman, S.; Axler, S.; Gehring, F. *Advanced Linear Algebra*; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3.
14. Dines, L.L. On Positive Solutions of a System of Linear Equations. *Ann. Math.* **1926**, *28*, 386–392. [CrossRef]
15. Schrijver, A. *Theory of Linear and Integer Programming*; John Wiley & Sons: Hoboken, NJ, USA, 1998.
16. Dantzig, G.B.; Thapa, M.N. *Linear Programming 1: Introduction*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2006.
17. Karmarkar, N. A new polynomial-time algorithm for linear programming. In Proceedings of the Sixteenth Annual ACM Symposium on Theory of Computing, Washington, DC, USA, 30 April–2 May 1984; pp. 302–311.
18. Khachiyan, L.G. Polynomial algorithms in linear programming. *USSR Comput. Math. Math. Phys.* **1980**, *20*, 53–72. [CrossRef]
19. Potra, F.A.; Wright, S.J. Interior-point methods. *J. Comput. Appl. Math.* **2000**, *124*, 281–302. [CrossRef]
20. Wright, M. The interior-point revolution in optimization: History, recent developments, and lasting consequences. *Bull. Am. Math. Soc.* **2005**, *42*, 39–56. [CrossRef]
21. Dantzig, G. *Linear Programming and Extensions*; Princeton University Press: Princeton, NJ, USA, 2016.
22. Dantzig, G.B.; Thapa, M.N. *Linear Programming 2: Theory and Extensions*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2006.
23. Murty, K.G. *Linear Programming*; Springer: Berlin/Heidelberg, Germany, 1983.
24. Bro, R.; De Jong, S. A fast non-negativity-constrained least squares algorithm. *J. Chemom. J. Chemom. Soc.* **1997**, *11*, 393–401. [CrossRef]
25. Gill, P.E.; Murray, W.; Wright, M.H. *Practical Optimization*; SIAM: Philadelphia, PA, USA, 2019.
26. Van Benthem, M.H.; Keenan, M.R. Fast algorithm for the solution of large-scale non-negativity-constrained least squares problems. *J. Chemom. J. Chemom. Soc.* **2004**, *18*, 441–450. [CrossRef]
27. Franc, V.; Hlaváč, V.; Navara, M. Sequential coordinate-wise algorithm for the non-negative least squares problem. In Proceedings of the International Conference on Computer Analysis of Images and Patterns, Versailles, France, 5–8 September 2005; Springer: Berlin/Heidelberg, Germany, 2005; pp. 407–414.
28. Kim, D.; Sra, S.; Dhillon, I.S. *A New Projected Quasi-Newton Approach for the Nonnegative Least Squares Problem*; Citeseer: Princeton, NJ, USA, 2006.
29. Bellavia, S.; Macconi, M.; Morini, B. An interior point Newton-like method for non-negative least-squares problems with degenerate solution. *Numer. Linear Algebra Appl.* **2006**, *13*, 825–846. [CrossRef]

30. Cantarella, J.; Piatek, M. Tsnnls: A solver for large sparse least squares problems with non-negative variables. *arXiv* **2004**, arXiv:cs/0408029.
31. Chen, D.; Plemmons, R.J. Nonnegativity constraints in numerical analysis. In *The Birth of Numerical Analysis*; World Scientific: Singapore, 2010; pp. 109–139.
32. Portugal, L.F.; Judice, J.J.; Vicente, L.N. A comparison of block pivoting and interior-point algorithms for linear least squares problems with nonnegative variables. *Math. Comput.* **1994**, *63*, 625–643. [CrossRef]
33. Lawson, C.; Hanson, R. *Solving Least-Squares Problems*; Prentice-Hall: Upper Saddle River, NJ, USA, 1974; Chapter 23.
34. Zhang, Y.Y.; Yu, H.Y.; Zhang, J.K.; Wang, J.L. Reliable MIMO Optical Wireless Communications Through Super-Rectangular Cover. *arXiv* **2016**, arXiv:1607.04206.
35. Chu, X. Hyper-Rectangle Cover Theory and Its Applications. Ph.D. Thesis, McMaster University, Hamilton, ON, Canada, 2022.
36. Grünbaum, B.; Klee, V.; Perles, M.A.; Shephard, G.C. *Convex Polytopes*; Springer: Berlin/Heidelberg, Germany, 1967; Volume 16.
37. Roman, S. Positive Solutions to Linear Systems: Convexity and Separation. In *Advanced Linear Algebra*; Springer: New York, NY, USA, 2005; pp. 395–408. [CrossRef]
38. Chu, M.; Plemmons, R. Nonnegative matrix factorization and applications. *Bull. Int. Linear Algebra Soc.* **2005**, *34*, 26.
39. Petrou, M.M.; Petrou, C. *Image Processing: The Fundamentals*; John Wiley & Sons: Chichester, West Sussex, UK, 2010.
40. Fu, X.; Huang, K.; Sidiropoulos, N.D.; Ma, W.K. Nonnegative Matrix Factorization for Signal and Data Analytics: Identifiability, Algorithms, and Applications. *IEEE Signal Process. Mag.* **2019**, *36*, 59–80. [CrossRef]
41. Difonzo, F.V. A note on attractivity for the intersection of two discontinuity manifolds. *Opusc. Math.* **2020**, *40*, 685–702. [CrossRef]
42. Dieci, L.; Difonzo, F. The moments sliding vector field on the intersection of two manifolds. *J. Dyn. Differ. Equ.* **2017**, *29*, 169–201. [CrossRef]
43. Trefethen, L.N.; Bau, D., III. *Numerical Linear Algebra*; SIAM: Philadelphia, PA, USA, 1997; Volume 50.
44. Klee, V.; Minty, G.J. How good is the simplex algorithm. *Inequalities* **1972**, *3*, 159–175.
45. Lawson, C.L.; Hanson, R.J. *Solving Least Squares Problems*; SIAM: Philadelphia, PA, USA, 1995.
46. Plemmons, R.J. M-matrix characterizations. I—nonsingular M-matrices. *Linear Algebra Its Appl.* **1977**, *18*, 175–188. [CrossRef]
47. Björck, Å. *Numerical Methods for Least Squares Problems*; SIAM: Philadelphia, PA, USA, 1996.