HYPER-RECTANGLE COVER THEORY AND ITS APPLICATIONS

HYPER-RECTANGLE COVER THEORY AND ITS APPLICATIONS

BY

XIAOXUAN CHU, B.S.

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Doctor of Philosophy (2022) McMaster University

(Electrical & Computer Engineering) Hamilton, Ontario, Canada

TITLE:	Hyper-Rectangle Cover Theory and Its Applications
AUTHOR:	Xiaoxuan Chu B.S. (Mathematics and Statistics), Zhengzhou University, Zhengzhou, China
SUPERVISOR:	Kon Max Wong Jun Chen Jiankang Zhang(Dec.)
NUMBER OF PAGES:	xv, 168

Abstract

In this thesis, we propose a novel *hyper-rectangle cover theory* which provides a new approach to analyzing mathematical problems with nonnegativity constraints on variables. In this theory, two fundamental concepts, *cover order and cover length*, are introduced and studied in details.

In the same manner as determining the rank of a matrix, we construct a specific échelon form of the matrix to obtain the cover order of a given matrix efficiently and effectively. We discuss various structures of the échelon form for some special cases in detail. Based on the structure and properties of the constructed échelon form, the concepts of *non-negatively linear independence* and *non-negatively linear dependence* are developed. 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 *nonnegativity constraints* on variables for both *homogeneous* and *non-homogeneous* cases. In addition, we apply the cover theory to analyze some typical problems in linear algebra and optimization with nonnegativity constraints on variables, including *linear programming problems* and *non-negative least squares* (NNLS) problems. For linear programming problem, we study the three possible behaviors of the solutions for it through hyper-rectangle cover theory, and show that a series of feasible solutions for the problem with the zero-cover échelon form structure.

On the other hand, we develop a method to obtain the cover length of the covered variable. In the 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, linear equations system, nonnegativity constraints, non-negative least squares, linear programming

To my family, for their love and support.

Acknowledgements

I am grateful to have the opportunity of working with all the people during the past four years.

First and foremost, I would like to express my gratitude to my supervisor, Jian-Kang Zhang (Dec), who was an academically rigorous, responsible and respectable scholar. Whenever he was engaged in research, he consistently came up with novel ideas and I had always been fascinated by his well-equipped mathematical abilities. Having him as a supervisor influenced me in a number of ways, not only in terms of academic guidance and assistance, but also for his friendly and considerate manner with people. I feel fortunate and proud to have been his student. If not for his encouragement and support, I would not have gained the opportunity to study as a Ph.D. student in Canada, experience the magic of mathematics and communications, and I would have lived a very different life. He treated me like his own children, and showed tremendous care and concern in my life, but he was also extremely strict when it came to my studies. His sudden passing away comes as a shock to me, and I deeply regret the loss of such a great fatherly supervisor. I will do my best to carry forward the knowledge that I have gained from him and as a person in the future. May he rest in peace!

Furthermore, I would like to extend my gratitude to Dr. Kon Max Wong, for his continuous encouragement, support and stimulating discussions. He provided me with support when I needed it most and after Dr. Zhang left, took me on as his student, providing me with guidance and assistance. In addition, he is always willing to assist me when I encounter difficulties, not only in my work but also in my personal life, with his kindness and generosity. During the last period of my Ph.D. studies, he helped a lot in revising my paper and the thesis for my dissertation. Through his assistance and guidance, I am able to develop better writing skills and express my thoughts clearly.

My sincere appreciation also goes to my supervisor, Dr. Jun Chen, for his guidance and support throughout my research studies. His inventive and creative ideas always make me feel so enlightened and inspired. It is always a pleasure to discuss with Dr. Chen and he acts both as a knowledgeable mentor and a close friend.

Great appreciations are also expressed to my committee members Dr. Sorina Dumitrescu and Dr. Rong Zheng, for their valuable guidance and comments on my research works.

Moreover, I appreciate the help and moral support from all my colleagues in the Department of Electrical and Computer Engineering.

Last but not least, I would like to thank my mother for her endless love and support in my life. My appreciation also goes out to my husband, Jingfeng Liu, without his tremendous understanding and encouragement in the past four years, it would be hard for me to complete my study.

Contents

List of Figures

List of Tables

Notation and Abbreviations

Notation

e.g., x	column vectors are denoted by lowercase boldface characters
e.g., A	matrices are denoted by uppercase boldface characters
e.g., $\mathcal{S}$	sets contain elements or members
e.g., $\lvert\mathcal{S}\rvert$	the number of members of set $\mathcal{S}$
$\mathbf{A}^T$	transpose of matrix $\mathbf{A}$
$\mathbf{A}^{-1}$	inverse of matrix $\mathbf{A}$
$\mathbf{A}_{ij}$	the (i, j)-th element of matrix $\mathbf{A}$
$\mathbf{det(A)}$	the determinant of matrix $\mathbf{A}$
$\lvert\mathbf{A}\rvert$	the determinant of matrix $\mathbf{A}$
$\parallel \mathbf{A} \parallel_2$	the Euclidean norm of matrix $\mathbf{A}$
$\mathbf{I}_{N\times N}$	$N \times N$ identity matrix

$\mathbb{R}$	field of real numbers
$\mathbb{R}_{+}^{N}$	the set of all the $N \times 1$ vectors with all entries being nonnegative
$\mathbb{R}_{++}^{N}$	the set of all the $N \times 1$ vectors with all entries being positive
e.g., $\mathbf{x} \geq \mathbf{0}$	non-negative column vectors
e.g., $\mathbf{x} > \mathbf{0}$	positive column vectors
e.g., $\mathbf{A} \geq \mathbf{0}$	non-negative matrices with all entries are non-negative
e.g., $\mathbf{A} > \mathbf{0}$	positive matrices with all entries are positive
x_i	the i-th element of the column vector $\mathbf{x}$
$\bar{\mathbf{A}}_{ij}$	sub-matrix formed by deleting i-th row and j-th column from $\mathbf{A}$
$\mathbf{rank(A)}$	rank of matrix $\mathbf{A}$
$\mathbf{A}^{\dagger}$	pseudoinverse of matrix $\mathbf{A}$
$\mathbf{A} \succ \mathbf{0}$	the matrix $\mathbf{A}$ is a positive definite matrix
$\mathbf{A} \succeq \mathbf{0}$	the matrix $\mathbf{A}$ is a positive semi-definite matrix
$\emptyset$	empty set

Abbreviations

PSD	Positive Semi-Definite Matrix
PD	Positive Definite Matrix
NNLS	Non-Negative Least Squares
LP	Linear Programming
FNNLS	Fast Non-Negative Least Squares
PQN-NNLS	Projective Quasi-Newton Non-Negative Least Squares
SCA	Sequential Coordinate-Wise Algorithm
NMF	Non-Negative Matrix Factorization

Chapter 1

Introduction

1.1 Motivation

The problems with nonnegativity constraints on variables play a prominent role in engineering, physics, chemistry, computer science, and economics. In linear algebra, nonnegativity constraints are frequently encountered in the system of linear equations, non-negative least squares problems, and linear programming problems (40). However, most existing methods for solving linear equations system with nonnegativity constraints are based on the analysis of their dual problems. Also, most existing algorithms for solving NNLS problems are based on numerical analysis rather than the matrix itself. We are thus motivated to develop a novel mathematical tool that can be applied to analyze the problems with nonnegativity constraints on variables in a more direct way.

1.2 Previous Work and Current Challenge

The system of linear equations is a fundamental part of linear algebra, which is widely used throughout modern mathematics. There are various mature methods that can be used to perform the analysis of the solution of linear equations system without nonnegativity constraints, such as investigating the rank of the corresponding matrices (18; 29; 51). However, the addition of nonnegativity constraints for the variables makes the analysis of the solutions to the linear equations and the method to determine whether it has non-zero solution directly more complex (19). The analysis of the existence of non-negative solution is mainly based on Farkas' Lemma (52). In terms of uniqueness, there is no direct characterization in the general case. It should be noted that in the analysis of non-negative solutions to the system of homogeneous or non-homogeneous linear equations are mostly carried out by investigating other associated problems rather than addressing the problem in a direct way. A novel approach is therefore needed for the analysis of the system of linear equations with nonnegativity constraints on variables directly.

Linear programming problems arise in many applications (16). Often, problems can be reformulated as linear programs both in theory and in practice so that fast algorithms can be utilized. Now, in a linear programming problem with nonnegativity constraints on variables, there is an objective function which is linearly dependent on a number of independent variables and those variables must also satisfy the non-negativity condition. Dantzig invented the *simplex method* in 1947, which was the first efficient method for solving linear programming problems and has been widely accepted as a computational tool (15). The simplex method is quite efficient in practice, and the global optimal solution can be guaranteed as long as certain precautions are taken to prevent cycling (17; 43). Klee and Minty, however, found a family of linear programming problems for which the simplex method takes a number of steps which grows exponentially with the size of the problem (35). For some time, it was uncertain whether linear programming could be solved in polynomial time until 1979 when Leonid Khachiyan demonstrated that linear programming can be solved in polynomial time (33). Narendra Karmarkar introduced the interior-point method for solving linear programming problems in 1984, which represents a significant theoretical and practical advancement in the field (31). The interior-point method was developed as a result of the desire to develop algorithms with more solid theoretical foundations than the simplex method. Despite some similarities between those two strategies, the interior- point methods involve relatively expensive (in terms of computing) iterations that quickly arrive at a solution, whereas the simplex method typically involves many more inexpensive iterations (48; 61). Geometrically, interior-point methods approach a solution from the interior or the exterior of a feasible region, but never from the boundary. Considering simplex method, the procedure of which involves moving one feasible solution to another and the value of the objective function will improve at each step. In this thesis, cover theory will be applied to analyze the possibilities of the solutions towards the LP problem, and for some matrix that contains the certain structures, the optimal solution can be obtained in a direct way by using the properties of cover order.

In mathematical optimization, non-negative least squares (NNLS) problems have been in existence for a long time, and are a type of least squares problem with nonnegativity constraints widely encountered in science and engineering. Various methods have been proposed to solve NNLS problems. Normally these methods can be divided into three classes: *active set methods* (24), *iterative approaches* and other methods (13). The first technique to solve it is proposed by Lawson and Hanson in (38). It is a typical example of an active set method and the corresponding algorithm is named as **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 optimum solution. The immediate followup to this work is called Fast NNLS (FNNLS), which is developed by Bro and Jong (9). This method speeds up the basic NNLS by avoiding unnecessary recomputations. By appropriately rearranging calculations to achieve further speedups in the presence of multiple observation vectors, a variant of FNNLS, called fast combinatorial NNLS (56) is proposed. In contrast to active set methods, the iterative approaches enable one to incorporate multiple active constraints at each iteration. The Projective Quasi-Newton NNLS (PQN-NNLS) algorithm (34) is a representative in this category, This method is based on Newton iteration and by using non-diagonal gradient scaling matrix to approximate the hessian matrix at each iteration, the rate of convergence accelerates efficiently. The novel sequential coordinate-wise (SCA) algorithm (21) is another example of an iterative method for solving NNLS problem, the idea of which is to optimize in each iteration with respect to a single coordinate while the remaining coordinates are fixed. Apart from these two major categories, interior point Newton-like method (3) generates an infinite sequence of strictly feasible points that converge to the solution and is known to be competitive with active set methods for medium and large problems. Finally, the principal block pivoting method (12; 47) is competitive for large and sparse NNLS problem. Since most existing algorithms for solving NNLS problem are based on numerical analysis, we are motivated to derive a method to solve it from the matrix perspective by applying the techniques we developed in cover theory, specifically, by investigating the structure of matrix itself to obtain the closed-form optimal value of NNLS problem.

1.3 Contributions and Thesis Organization

All the aforementioned factors motivate us to propose a novel theory for the deep analysis of the problems that arise in the linear algebra and optimization with the nonnegativity constraints on the variables. The main contributions in this thesis are summarized as follows:

1) First, we establish the basic principles of the novel hyper-rectangle cover theory which will be utilized as a basis for the discussion that follows. After the verification of whether a variable is covered or not, we obtain the result of the equivalence between the full cover matrix and the existence condition of non-zero solution towards the system of homogeneous linear equations with nonnegativity constraints. In addition, we establish conditions for guaranteeing the existence of non-negative solutions to the system of non-homogeneous linear equations with nonnegativity constraints. Moreover, by using the cover theory, we are able to find the necessary and sufficient conditions under which a system of non-homogeneous linear equations with nonnegativity constraints has a unique solution. It is noted that these conditions are based on the matrix itself, in other words, we are able to analyze the solution of the system of linear equations with nonnegativity constraints from analysing the matrix itself rather than by investigating other associated problems.

2) Second, we present the related properties of cover theory which can be used to determine the cover order and we also propose a specific échelon form of the matrix.

Based on this échelon form, an efficient and effective method has been developed to determine the cover order for any given matrix. The special structure of low-rank matrices and matrices with some special form have been investigated. We also analyze the structure of zero-cover matrices in depth, which can be useful to obtain the feasible solution for the system of linear equations with nonnegativity constraints on solutions and the linear programs. In addition, we develop the concepts about non-negatively linear independence and non-negatively linear dependence, with which we could gain a deeper insight about the linear equations system with nonnegativity constraints on solutions. Additionally, we obtain the dual property between cover and uncover by using the definitions and properties of generalized inverse of the matrix. We also introduce an inner hyper-rectangle concept which has a strong relationship with the properties of zero-cover.

3) Furthermore, on the basis of the specific échelon form and the corresponding results on the system of linear equations with nonnegativity constraints on solutions, we can verify whether or not the feasibility set of the linear programming problem is empty. Then the various possibilities of the solutions and the optimal values of the linear programs, to be more specific, it has optimal bounded solution, the linear program is feasible but unbounded, or it has infinite unbounded optimal solution , corresponding respectively to the following scenarios: full cover, zero cover or the cover order in between, have been analyzed in detail. Furthermore, with the special structure of zero-cover part in the échelon form of the matrix, we are able to arrive at a series of feasible solutions to the linear programs. We also compare the cover method and the simplex method in solving LP problem and at the mean time, we apply the proposed method to efficiently and effectively solve the Klee-Minty cube

problem, which will require exponential time to solve if the simplex method is utilized.

4) Finally, we develop a method to determine the cover length of the covered variables. We also establish the strong connection between solving NNLS problems and finding the cover length. Based on this relationship, the NNLS problem can be reconstructed as the cover length determination problem from which we are able to obtain the analytical optimal value of the NNLS problem directly by analyzing the matrix itself. For some certain types of matrices, such as the M-matrix, the closed-form optimal value of the objective function of NNLS problem can be determined in a more direct manner. In order to gain a deeper understanding of cover length determination method, we also compare our proposed method to the commonly used method, active-set method, in solving NNLS problem.

The rest of this thesis is organized as follows:

- **Chapter 2:** The basic related definitions and properties of novel hyper-rectangle cover theory.
- **Chapter 3:** The character of full-cover and the analysis of the system of linear equations with nonnegativity constraints on solutions based on the cover theory.
- **Chapter 4:** The details of the proposed échelon form to determine the cover order and the structure of zero-cover matrix.
- **Chapter 5:** The details of the new systematic procedure for solving linear programming problems by applying cover theory.
- **Chapter 6:** The method to determine the cover length of the covered variable and utilize this method to attain the analytical optimal value of NNLS problem.

- **Chapter 7:** The conclusion of this thesis and the a discussion of the possible future work.

The logical connections among the chapters are shown in the Figure 1.1. We start with the basic concepts of hyper-rectangle cover theory in Chapter 2, then we apply the cover theory in analyzing the linear equations system with nonnegativity constraints on solutions in Chapter 3. Then we can either go to Chapter 4 for the investigation of the properties of cover order and go further to its applications in solving linear programming problem in Chapter 5, or we can go directly to Chapter 6 to examine the properties of cover length and its applications in the analysis of non-negative least squares problem. Finally, Chapter 7 includes the conclusion of above and some possible future work.

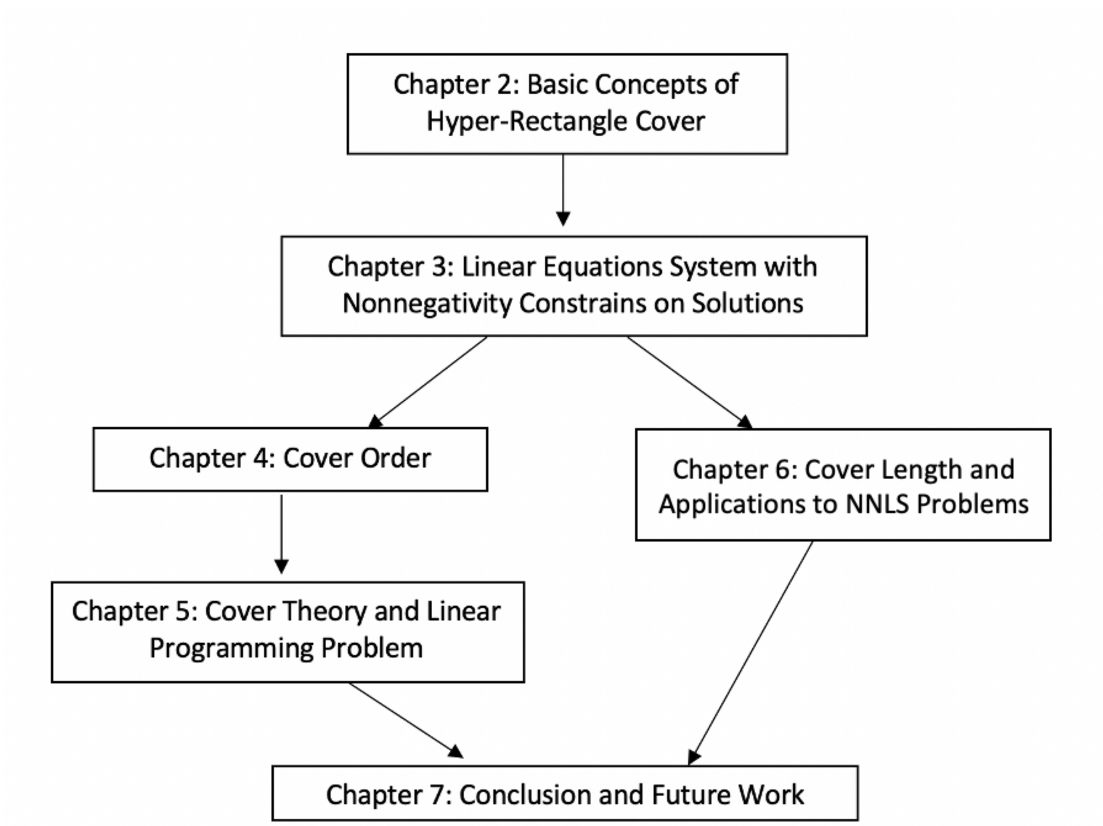

Figure 1.1: Logical connections among the chapters.

Chapter 2

Basic Concepts of Hyper-Rectangle Cover

In this chapter, for an $M \times N$ real matrix $\mathbf{A}$, we formally give the definition of hyper-rectangle cover (63). Two basic concepts are introduced: 1) cover order, which is defined by the maximal side number of the hyper-rectangle that covers the feasible domain determined by a quadratic form smaller than any given positive constant in the non-negative orthant; and 2) cover length, which is defined as the side length of the smallest hyper-rectangle that covers above feasible domain.

2.1 Definition of Hyper-Rectangle Cover

Definition 2.1.1. *An $M \times N$ real matrix $\mathbf{A}$ is said to be r-covered ($r \leq N$) by a hyper-rectangle, if, for any given positive real-valued number $\tau > 0$, the domain determined by $\{\mathbf{x} : \mathbf{x} \in \mathbb{R}_+^N, \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2\}$ is located inside this hyper-rectangle*

$\{\mathbf{x} : 0 \leq x_{n_k} \leq c_{n_k}\tau, k = 1, 2, \cdots, r\}$, *i.e.*,

$$\{\mathbf{x} : \mathbf{x} \in \mathbb{R}_+^N, \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2\}$$
$$\subseteq \{\mathbf{x} : \mathbf{x} \in \mathbb{R}_+^N, 0 \leq x_{n_k} \leq c_{n_k}\tau, k = 1, 2, \cdots, r\}, \quad (2.1.1)$$

where all c_{n_k} are constants independent of τ. In addition,

1. *the maximum r is named the cover order of $\mathbf{A}$ and denoted by $R_c(\mathbf{A})$.*
2. *When $R_c(\mathbf{A}) = N$, $\mathbf{A}$ has a full-cover.*
3. *When $R_c(\mathbf{A}) = 0$, $\mathbf{A}$ has a zero-cover.*
4. *For a fixed n_k satisfying $1 \leq k \leq R_c(\mathbf{A})$, the minimum constant c_{n_k} is named the n_k-th cover length.*

■

Those bounded elements in the variable $\mathbf{x}$ in the second set of Eq. (2.1.1) are covered variables, and the corresponding column vectors in $\mathbf{A}$ are covered columns. The above definition provides us with an intuitive method to determine the cover order of any given matrix and the cover length of the corresponding covered variable. For any matrix $\mathbf{A} \in \mathbb{R}^{M\times N}$ and non-negative vector $\mathbf{x} \in \mathbb{R}_+^N$, the cover order of $\mathbf{A}$ is the maximum number of bounded variable x_i in $\mathbf{x}$ when the inequality $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2$ holds, where $i \in \{1, 2, \cdots, N\}$. The cover length is the maximum value that the bounded variable can arrive.

For the purposes of better understanding the discussions above, the following non-trivial examples are provided to illustrate the definitions of cover order and cover length.

Example 2.1.1. *Consider*

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \end{pmatrix}$$

and thus, we have

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

In this case, we have

$$\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = (x_1 + x_2)^2 \leq \tau^2,$$

where τ *is any given positive real-valued number. The domain* $\{\mathbf{x} \in \mathbb{R}_+^2 : \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2\}$ *is shown in Figure 2.1. As illustrated by Figure 2.1, it can be seen that the domain* $\{\mathbf{x} \in \mathbb{R}_+^2 : \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2\}$ *is triangle-shaped with three vertices being* $(0, 0), (1, 0)$ *and* $(0, 1)$ *and can be covered by a square with four vertices determined by* $(0, 0), (1, 0), (0, 1),$ *and* $(1, 1)$. *Hence, for any* $\tau > 0$,

$$\begin{aligned} &\{\mathbf{x} \in \mathbb{R}_+^2 : \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2\} \\ &\subseteq \{\mathbf{x} \in \mathbb{R}_+^2 : 0 \leq x_1, x_2 \leq \tau\} \end{aligned}$$

Therefore, x_1 *and* x_2 *are all covered and* $R_c = 2$, *then,* $\mathbf{A}$ *has a full-cover. The cover length of* x_1 *and* x_2 *are both 1.* ■

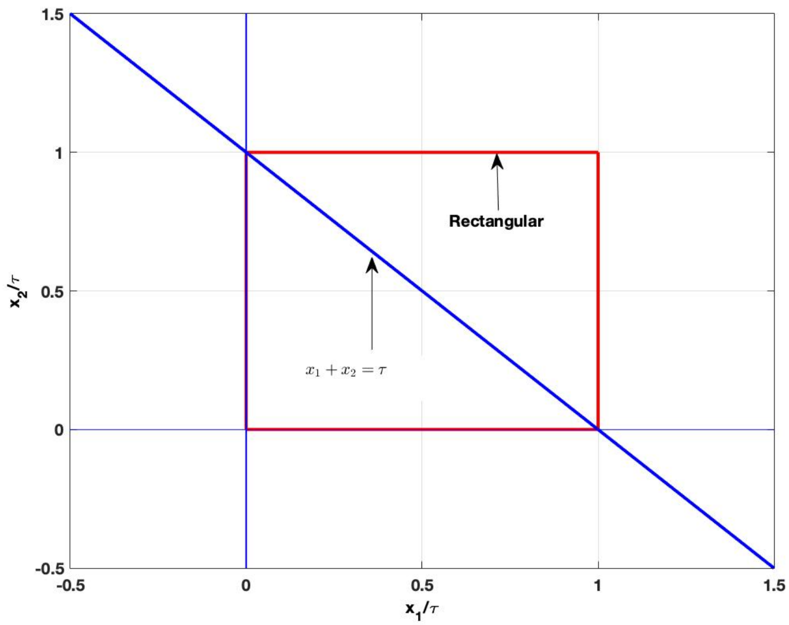

Figure 2.1: The full-cover matrix in Example 2.1.1

Example 2.1.2. *Consider*

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \end{pmatrix}$$

and then we have

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

For this matrix, we have the domain

$$\begin{aligned} &\{\mathbf{x} \in \mathbb{R}_+^2 : \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2\} \\ \subseteq\ &\{\mathbf{x} \in \mathbb{R}_+^2 : 0 \leq x_2 \leq \tau, x_1 \geq 0\} \end{aligned}$$

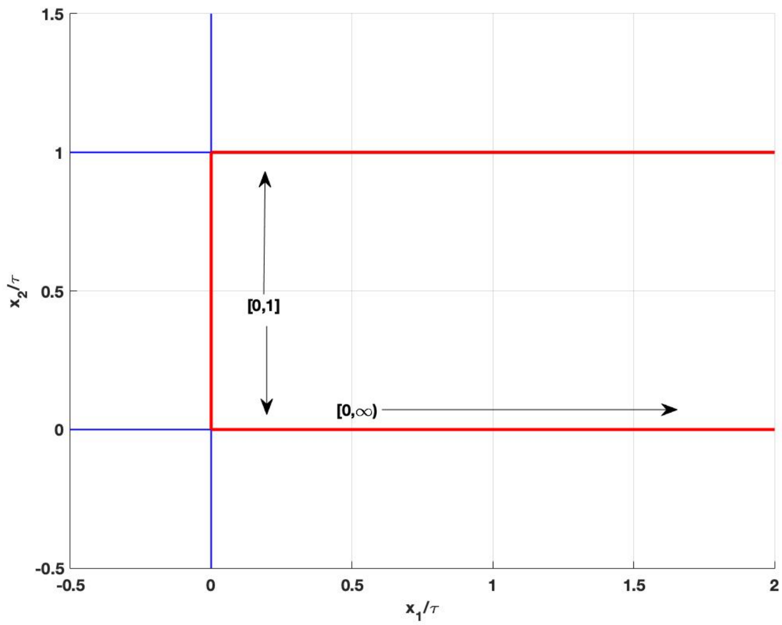

Figure 2.2: The one-cover matrix in Example 2.1.2

This domain is shown in Figure 2.2 and it is shown to be unbounded in the x_1*-axis direction. In other words, only one dimension is covered. Thus, we obtain that* $R_c = 1$ *and the cover length of the covered variable* x_2 *is* $c_2 = 1$*.* ■

Example 2.1.3. *Consider*

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \end{pmatrix}$$

and we have

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},$$

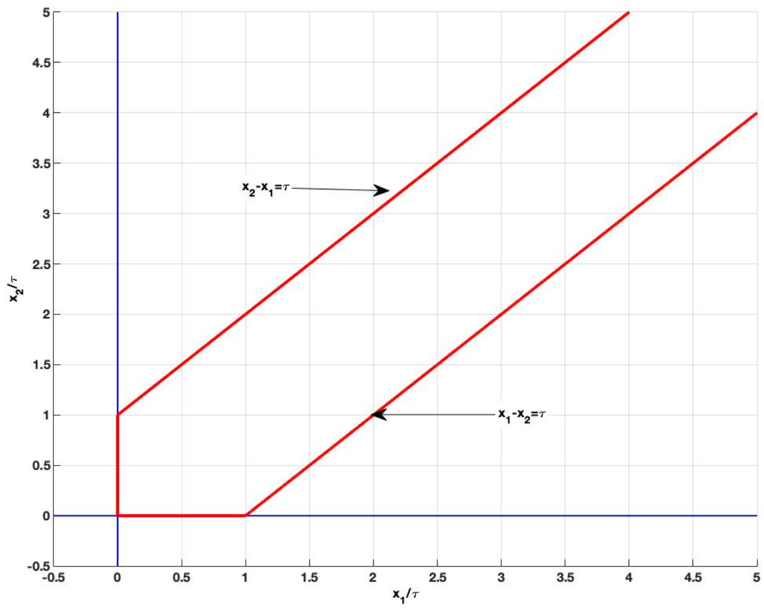

Figure 2.3: The zero-cover matrix in Example 2.1.3

for this case, the feasible set:

$$\{\mathbf{x} \in \mathbb{R}^2_+ : \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = (x_1 - x_2)^2 \leq \tau^2\}$$

is shown in Figure 2.3 and is open and unbounded with respect to both x_1 *and* x_2. *Hence,* $\mathbf{A}$ *has zero-cover, i.e.,* $R_c(\mathbf{A}) = 0$. ■

Example 2.1.4. *Suppose*

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

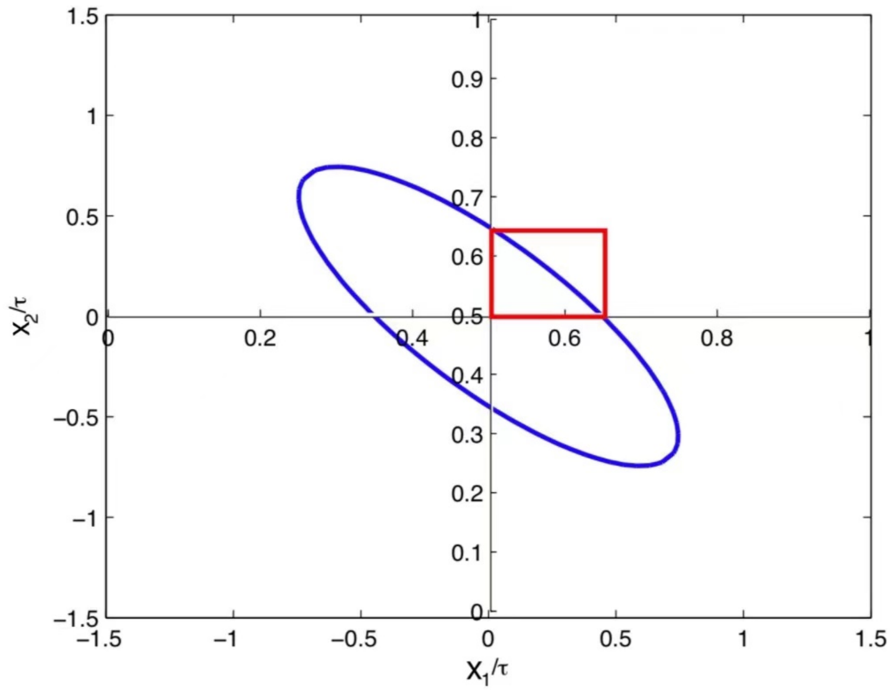

Figure 2.4: The 2×2 full cover matrix in Example 2.1.4

and we have the following matrix

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}.$$

For this matrix, the set $\{\mathbf{x} \in \mathbb{R}_+^2 : \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2\}$ *is an ellipse in the whole plane, which is shown in Figure 2.4 and the part which is located in the positive domain is fully covered by a rectangle. Thus* x_1 *and* x_2 *are both covered in this example and* $\mathbf{A}$ *has a full-cover.* ■

From the above four examples, we can notice that the cover order R_c of the matrix $\mathbf{A}$ represents the maximum dimension of the hyper-rectangle that covers the domain determined by the set $\{\mathbf{x} : \mathbf{x} \in \mathbb{R}_+^N, \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2\}$ and the cover length c_i of the covered variable x_i in $\mathbf{x}$ represent the side length of the minimum hyper-rectangle that covers above domain.

2.2 Conclusion

Among all the hyper-rectangles that cover the feasible domain determined by a quadratic form smaller than any given positive constant in the non-negative orthant, the cover order is the number of dimensions of the hyper-rectangle that covers the domain which has the maximum number of finitely lengthy sides, (for example, in Figure 2.2, the "hyper-rectangle" only has finite side length in x_2-dimension, thus the cover order is 1). and the cover length is the side length of hyper-rectangle that covers the domain which has the minimum length of each finite-length side.

Chapter 3

Linear Equations System with Nonnegativity Constraints on Solutions

In this chapter, we propose the condition under which the variable and the corresponding column vector in the matrix are covered. This result enables us to arrive at discoveries in the analysis of the solutions for the linear equations system with nonnegativity constraints. We reveal the relationship between full cover and the unique solution of the system of homogeneous linear equations with non-negative constraints, which further provides us the condition for the existence of non-zero solution for such a system. In terms of the system of non-homogeneous linear equations with nonnegativity constraints, we also derive the necessary and sufficient condition with regard to the existence and the uniqueness of the solutions.

With the aid of these newly developed results, the analysis of the non-negative solutions to the system of liner equations would be more straightforward. One of the major advantages of these findings lie in the fact that these conditions are based on the matrix itself, in other words, we are able to analyze the solution of the system of linear equations with nonnegativity constraints on solutions from the matrix itself rather than by investigating other associated problems, like transforming the original problem into a non-negative least squares problem or a linear programming problem.

3.1 System of Homogeneous Linear Equations with Nonnegativity Constraints on Solutions

Let $\mathbf{A}$ be an $M \times N$ real matrix and $\mathbf{x}$ is a real column vector with N elements. We present an important result in cover theory, allowing us to determine whether the i-th column vector $\mathbf{a}_i$ in $\mathbf{A}$ or the corresponding i-th variable x_i of $\mathbf{x}$ in $\mathbf{Ax}$ is covered or not, where $i \in \{1, \cdots, N\}$. Furthermore, considering the system of homogeneous linear equations with nonnegativity constraints on the solution: $\mathbf{Ax} = \mathbf{0}$, where $\mathbf{x} \geq \mathbf{0}$, new approaches of investigating the existence as well as the uniqueness of solutions to such a system are proposed.

Theorem 3.1.1. *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. To begin with, let us prove the necessary condition, i.e., by assuming that the i-th column vector in the matrix $\mathbf{A}$ or the i-th variable x_i in the vector $\mathbf{x}$ is covered in $\mathbf{Ax}$, we need to prove that for any non-negative real vector $\mathbf{x} \in \mathbb{R}_+^N$ with $x_i > 0$, the product $\mathbf{Ax}$ is non-zero. Suppose that this statement was not true. Then, we would have that there exists $\mathbf{x}_0 \in \mathbb{R}_+^N$ with $x_{0,i} > 0$ such that the equality $\mathbf{Ax}_0 = \mathbf{0}$

holds. As a consequence, for any positive real number $p > 0$, we would also have $\mathbf{A}(p\mathbf{x}_0) = \mathbf{0}$, which implies $px_{0,i}$ is not bounded and for any given positive real-valued $\tau > 0$, we have:

$$0 = (p\mathbf{x}_0)^T \mathbf{A}^T \mathbf{A}(p\mathbf{x}_0) \leq \tau \tag{3.1.1}$$

This contradicts with the assumption that x_i is covered in $\mathbf{Ax}$. Therefore, the necessary condition is true.

Now, let us consider the proof of the sufficient condition. Notice that for non-negative real vector $\mathbf{x}$ with $x_i > 0$, the quadratic form $\mathbf{x}^T\mathbf{A}^T\mathbf{Ax}$ can be rewritten as follows:

$$\mathbf{x}^T\mathbf{A}^T\mathbf{Ax} = \|\bar{\mathbf{A}}_i\bar{\mathbf{x}}_i + \mathbf{a}_i x_i\|^2 = x_i^2\|\bar{\mathbf{A}}_i\mathbf{u} + \mathbf{a}_i\|^2 \tag{3.1.2}$$

where $\bar{\mathbf{x}}_i$ is the $(N-1) \times 1$ vector formed by deleting the i-th element x_i from $\mathbf{x}$, $\mathbf{u} = \bar{\mathbf{x}}_i/x_i$, $\bar{\mathbf{A}}_i$ is the $M \times (N-1)$ sub-matrix formed by deleting the i-th column from $\mathbf{A}$ and $\mathbf{u} \geq \mathbf{0}$. Considering the set $\{\bar{\mathbf{A}}_i\mathbf{u} : \mathbf{u} \in \mathbb{R}^{N-1}, \mathbf{u} \geq 0\}$. Since it is a closed convex hull, the minimum of $\|\bar{\mathbf{A}}_i\mathbf{u} + \mathbf{a}_i\|^2$ exists, i.e., there exists a $\mathbf{u}_0 \in \mathbb{R}^{N-1}$ and $\mathbf{u}_0 \geq \mathbf{0}$ such that for any $\mathbf{u} \in \mathbb{R}^{N-1}$ and $\mathbf{u} \geq \mathbf{0}$, we have

$$\|\bar{\mathbf{A}}_i\mathbf{u}_0 + \mathbf{a}_i\|^2 \leq \|\bar{\mathbf{A}}_i\mathbf{u} + \mathbf{a}_i\|^2. \tag{3.1.3}$$

In fact, $\|\bar{\mathbf{A}}_i\mathbf{u}_0 + \mathbf{a}_i\| \neq 0$. Otherwise, we can construct a non-negative real vector

$\mathbf{x}_0 \in \mathbb{R}^{N-1}$ in the following way, i.e., let:

$$\begin{aligned} x_{0,i} &= 1, \\ x_{0,k} &= u_{0,k}, \;\; k = 1, 2, \cdots, i-1, \\ x_{0,k} &= u_{0,k-1}, \;\; k = i+1, \cdots, N, \end{aligned} \tag{3.1.4}$$

where $x_{0,i}$ is the i-th element in $\mathbf{x}_0$ and $x_{0,k}$ is the k-th element in $\mathbf{x}_0$ while $u_{0,k}$ is the k-th element in $\mathbf{u}_0$, for $k = 1, \cdots, i-1, i+1, \cdots, N$. Then, if $\|\bar{\mathbf{A}}_i \mathbf{u}_0 + \mathbf{a}_i\| = 0$, we have:

$$\mathbf{A}\mathbf{x}_0 = \bar{\mathbf{A}}_i \mathbf{u}_0 + \mathbf{a}_i = \mathbf{0}, \tag{3.1.5}$$

which contradicts with the assumption that $\mathbf{A}\mathbf{x} \neq \mathbf{0}$ for any $\mathbf{x} \in \mathbb{R}_+^N$ with $x_i > 0$. Now for any given positive real-valued number $\tau > 0$, if we let $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau$, then, by Eq. (3.1.3), we have:

$$x_i^2 \|\bar{\mathbf{A}}_i \mathbf{u}_0 + \mathbf{a}_i\|^2 \leq \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2 \tag{3.1.6}$$

Hence, we obtain:

$$0 < x_i < \frac{\tau}{\|\bar{\mathbf{A}}_i \mathbf{u}_0 + \mathbf{a}_i\|} \tag{3.1.7}$$

i.e., x_i is covered in $\mathbf{A}\mathbf{x}$. This completes the proof of the sufficient condition. Thus the proof of Theorem 3.1.1 is finished. □

To provide some more insight into above theorem, we propose the following brief discussion. Given an $M \times N$ real matrix $\mathbf{A}$ and real column vector $\mathbf{x}$ with N elements in it, we consider the homogeneous linear equations system, $\mathbf{A}\mathbf{x} = \mathbf{0}$ with non-negative constraints $\mathbf{x} \geq \mathbf{0}$. Clearly, it is of interest to determine conditions that will guarantee the existence of non-zero solutions to this system. Analyzing the solutions of above homogeneous linear equations without nonnegativity constraints is quite straightforward (42). However, when we add the nonnegativity constraints of the solutions into the system, except transforming it into a related linear programming problem, the method to direct determine whether it has non-zero solution is not simple anyway (19; 50). It is noted that above theorem provides us with a method to solve the above problem and by Theorem 3.1.1 we are able to obtain the following results.

Corollary 3.1.1. *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 four statements are true:*

1. *A system of homogeneous linear equations:* $\mathbf{A}\mathbf{x} = \mathbf{0}$*, 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 have a full-cover.*

■

It must be noted that Theorem 3.1.1 is parallel to the first statement of Corollary 3.1.1, which provides us the necessary and sufficient condition for the existence of non-zero solutions for the homogeneous linear equations system with nonnegativity constraints on solutions. It is shown that this condition is direct and simple and it only relies on the matrix itself.

Example 3.1.1. *Following are examples that demonstrate how to determine whether a homogeneous linear equations system with nonnegativity constraints on solutions has nonzero solutions or not.*

1. *Consider* $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$*, according to the definition of cover order in Definition 2.1.1,* $\mathbf{A}$ *has a full cover, then by Theorem 3.1.1 and the first statement in Corollary 3.1.1, we can conclude that* $\mathbf{A}\mathbf{x} = \mathbf{0}$ *only has zero solution in* $\mathbb{R}^2_+$

2. *Consider* $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$*, we have* $R_c(\mathbf{A}) = 1$*. Since* $\mathbf{A}$ *does not have full cover,* $\mathbf{A}\mathbf{x} = \mathbf{0}$ *has non-zero solution in* $\mathbb{R}^2_+$*. We can notice that* x_2 *is an uncovered variable and it can be any non-negative value while* x_1 *is a covered variable and it must be zero.*

3. *Consider* $\mathbf{A} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$*,* $\mathbf{A}$ *has zero-cover. Thus* $\mathbf{A}\mathbf{x} = \mathbf{0}$ *has non-zero solution in the* $\mathbb{R}^2_+$*. Besides,* x_1 *and* x_2 *are both uncovered variables and they can be any non-negative value.*

■

In the following, we will give some algebraic conditions on full cover. For the sake of simplicity, we first introduce some notations.

For any $\mathbf{A} \in \mathbb{R}^{M\times N}$ and $\mathbf{x} \in \mathbb{R}^N_+$, let us first assume that $\bar{\mathbf{x}}_i$ is the $(N-1)\times 1$ vector attained by deleting i-th entry from $\mathbf{x}$, $\bar{\mathbf{A}}_{ii}$ is the $(N-1)\times(N-1)$ sub-matrix formed by deleting i-th row and i-th column from $\mathbf{A}^T\mathbf{A}$, a_{ii} is the entry in the i-th row and i-th column of $\mathbf{A}$, and $\bar{\mathbf{a}}_i$ is the $(N-1)\times 1$ vector generated by deleting i-th entry from the i-th column of $\mathbf{A}^T\mathbf{A}$. If there exist n negative entries in $\bar{\mathbf{a}}_i$, then, for presentation simplicity, we denote the indices of the negative entries of $\bar{\mathbf{a}}_i$ by $i_1^{(-)}, \cdots, i_n^{(-)}$, where $0 \le n \le N-1$. The discriminant of the following equation:

$$a_{ii}x_i^2 + 2x_i\bar{\mathbf{a}}_i^T\bar{\mathbf{x}}_i + \bar{\mathbf{x}}_i^T\bar{\mathbf{A}}_{ii}\bar{\mathbf{x}}_i = 0 \tag{3.1.8}$$

with respect to x_i is given by

$$\Delta_i \triangleq -4\bar{\mathbf{x}}_i^T(a_{ii}\bar{\mathbf{A}}_{ii} - \bar{\mathbf{a}}_i\bar{\mathbf{a}}_i^T)\bar{\mathbf{x}}_i \tag{3.1.9}$$

For a better understanding of the algebraic expression, we introduce the following definition of the Schur complement of a block matrix (62; 28).

Definition 3.1.1 (Schur Complement)**.** *Let* $\mathbf{M}$ *be an* $N\times N$ *matrix written as a* 2×2 *block matrix*

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}, \tag{3.1.10}$$

where $\mathbf{A}$ *is a* $p\times p$ *matrix and* $\mathbf{D}$ *is a* $q\times q$ *matrix, with* $N = p+q$*(so,* $\mathbf{B}$ *is a* $p\times q$ *matrix and* $\mathbf{C}$ *is a* $q\times p$ *matrix). Then:*

1. *If* $\mathbf{D}$ *is invertible, then the Schur complement of the block* $\mathbf{D}$ *of the matrix* $\mathbf{M}$ *is the* $p\times p$ *matrix defined by*

$$\mathbf{M}/\mathbf{D} := \mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C} \tag{3.1.11}$$

2. *If* $\mathbf{A}$ *is invertible, then the Schur complement of the block* $\mathbf{A}$ *of the matrix* $\mathbf{M}$ *is the* $q \times q$ *matrix defined by*

$$\mathbf{M}/\mathbf{A} := \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \tag{3.1.12}$$

■

Thus the Schur complement of the block $\bar{\mathbf{A}}_{ii}$ of the matrix $\mathbf{A}^T\mathbf{A}$ is given as

$$\mathbf{A}/a_{ii} = \bar{\mathbf{A}}_{ii} - \frac{\bar{\mathbf{a}}_i\bar{\mathbf{a}}_i^T}{a_{ii}} \tag{3.1.13}$$

Furthermore, if we assume that $\mathbf{M}$ in Definition 3.1.1 is symmetric, then with regard to the Schur complement we have the following result (57), where we use the usual notation, $\mathbf{M} \succ \mathbf{0}$ to say that $\mathbf{M}$ is positive definite and the notation $\mathbf{M} \succeq \mathbf{0}$ to say that $\mathbf{M}$ is positive semi-definite.

Lemma 3.1.1. *For any symmetric matrix* $\mathbf{M}$ *of the form*

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{pmatrix} \tag{3.1.14}$$

if $\mathbf{A}$ *is invertible then the following properties hold:*

1. $\mathbf{M} \succ \mathbf{0}$ *iff* $\mathbf{A} \succ \mathbf{0}$ *and* $\mathbf{C} - \mathbf{B}^\mathbf{T}\mathbf{A}^{-1}\mathbf{B} \succ \mathbf{0}$.

2. *If* $\mathbf{A} \succ \mathbf{0}$, *then* $\mathbf{M} \succeq \mathbf{0}$ *iff* $\mathbf{C} - \mathbf{B}^\mathbf{T}\mathbf{A}^{-1}\mathbf{B} \succ \mathbf{0}$.

■

The above lemma tells us that $a_{ii}\bar{\mathbf{A}}_{ii} - \bar{\mathbf{a}}_i\bar{\mathbf{a}}_i^T$ is a PSD matrix if $\mathbf{A}^T\mathbf{A}$ is PSD and $a_{ii} \neq 0$. Thus, we have:

$$\Delta_i \leq 0,\ \forall i = 1, 2,\ \cdots,\ N \tag{3.1.15}$$

Then, let us form an $n \times n$ sub-matrix by the rows and columns of $a_{ii}\bar{\mathbf{A}}_{ii} - \bar{\mathbf{a}}_i\bar{\mathbf{a}}_i^T$ indexed by $i_1^{(-)},\ \cdots,\ i_n^{(-)}$. To simplify the presentation, this sub-matrix is denoted by $\bar{\mathbf{C}}_i^{(-)}$ in the following discussion.

Theorem 3.1.2. *For non-zero PSD matrix $\mathbf{A}^T\mathbf{A} \in \mathbb{R}^{N\times N}$, the following four statements are true.*

1. *Suppose that there exists at least one i such that $\bar{\mathbf{a}}_i$ is nonnegative, that is,*

$$\{\bar{\mathbf{a}}_i : \bar{\mathbf{a}}_i \in \mathbb{R}_+^{N-1}, i = 1, 2,\ \cdots,\ N\} \neq \emptyset$$

   Then, $\mathbf{A}^T\mathbf{A}$ is of full-cover if and only if all the diagonal entries of $\mathbf{A}^T\mathbf{A}$ are nonzero and $\bar{\mathbf{A}}_{ii}$ is of full-cover for any i satisfying $\bar{\mathbf{a}}_i \in \mathbb{R}_+^{N-1}$.

2. *$\mathbf{A}^T\mathbf{A}$ is of full-cover if there exists i such that $a_{ii} > 0$ and $(a_{ii}\bar{\mathbf{A}}_{ii} - \bar{\mathbf{a}}_i\bar{\mathbf{a}}_i^T)$ is of full-cover.*

3. *$\mathbf{A}^T\mathbf{A}$ is of full-cover if all the principal sub-matrices in $\mathbf{A}^T\mathbf{A}$ are of full-cover.*

4. *Suppose that there exists $\bar{\mathbf{a}}_i$ having at least one negative entry, say,*

$$\{\bar{\mathbf{a}}_i : \bar{\mathbf{a}}_i \notin \mathbb{R}_+^{N-1}, i = 1, 2,\ \cdots,\ N\} \neq \emptyset$$

   Then, $\mathbf{A}^T\mathbf{A}$ is full-cover only if $\bar{\mathbf{C}}_i^{(-)}$ is full-cover for any i satisfying $\bar{\mathbf{a}}_i \notin \mathbb{R}_+^{N-1}$.

■

Proof. The proof of Theorem 3.1.2 is based on the following equality

$$\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = a_{ii}x_i^2 + 2x_i\bar{\mathbf{a}}_i^T\bar{\mathbf{x}}_i + \bar{\mathbf{x}}_i^T\bar{\mathbf{A}}_{ii}\bar{\mathbf{x}}_i \tag{3.1.16}$$

Proof of Statement 1): If there exists i such that $\bar{\mathbf{a}}_i$ is a non-negative vector, then, for any $\mathbf{x} \in \mathbb{R}_+^N$, we have:

$$a_{ii}x_i^2 \geq 0,\ 2x_i\bar{\mathbf{a}}_i^T\bar{\mathbf{x}}_i \geq 0,\ \bar{\mathbf{x}}_i^T\bar{\mathbf{A}}_{ii}\bar{\mathbf{x}}_i \geq 0.$$

Thus, $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = 0$ if and only if $a_{ii}x_i^2 = 0$, $2x_i\bar{\mathbf{a}}_i^T\bar{\mathbf{x}}_i = 0$ and $\bar{\mathbf{x}}_i^T\bar{\mathbf{A}}_{ii}\bar{\mathbf{x}}_i = 0$. Since $a_{ii} > 0$, $a_{ii}x_i^2 = 0$ if and only if $x_i = 0$. In addition, if $\bar{\mathbf{A}}_{ii}$ has full cover, then, Theorem 3.1.1 tells us that $\bar{\mathbf{x}}_i^T\bar{\mathbf{A}}_{ii}\bar{\mathbf{x}}_i = 0$ if and only if $\bar{\mathbf{x}}_i = \mathbf{0}$. In this case, $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = 0$ if and only if $\mathbf{x} = \mathbf{0}$. By Theorem 3.1.1, we can conclude that $\mathbf{A}^T\mathbf{A}$ is of full-cover and thus, the sufficiency proof of Statement 1) is complete.

To prove the necessity of Statement 1) by contradiction, we consider the following two possibilities.

1. Let us suppose that there exists one i such that $a_{ii} = 0$. Then, there exists a vector $\mathbf{x}_0$ with the i-th entry being one and the other $(N-1)$ entries being zeros such that $\mathbf{x}_0^T\mathbf{A}^T\mathbf{A}\mathbf{x}_0 = 0$. This observation tells us that $\mathbf{A}^T\mathbf{A}$ is not of full-cover by Theorem 3.1.1. Thus, the positiveness of all the diagonal entries of $\mathbf{A}^T\mathbf{A}$ is a necessary condition for $\mathbf{A}^T\mathbf{A}$ being of full-cover.

2. If $\{\mathbf{a}_i : \mathbf{a}_i \in \mathbf{R}_+^{N-1}\} \neq \emptyset$, then, we suppose that there exists i such that $\mathbf{a}_i \in \mathbf{R}_+^{N-1}$ and $\bar{\mathbf{A}}_{ii}$ is not of full-cover. Then, by Theorem 3.2.2, we can always find

an $(N-1)\times 1$ nonzero vector $\bar{\mathbf{x}} \in \mathbb{R}_+^{N-1}$ such that $\bar{\mathbf{x}}^T \bar{\mathbf{A}}_{ii}^T \bar{\mathbf{A}}_{ii}\bar{\mathbf{x}} = 0$. Then, we form an $N \times 1$ vector $\mathbf{x}$ by letting $x_i = 0$ and $\bar{\mathbf{x}}_i = \bar{\mathbf{x}}$. As a result, this nonzero vector $\mathbf{x} \in \mathbb{R}_+^N$ assures that

$$\begin{aligned}\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} &= a_{ii}x_i^2 + 2h_i\bar{\mathbf{a}}_i^T\bar{\mathbf{x}}_i + \bar{\mathbf{x}}_i^T\bar{\mathbf{A}}_{ii}\bar{\mathbf{x}}_i \\ &= \bar{\mathbf{x}}_i^T\bar{\mathbf{A}}_{ii}\bar{\mathbf{x}}_i = 0 \end{aligned} \tag{3.1.17}$$

which implies that $\mathbf{A}^T\mathbf{A}$ is not of full-cover by Theorem 3.1.1.

Thus, Statement 1) is necessary for $\mathbf{A}^T\mathbf{A}$ being full-cover. Then, the proof of Statement 1) is complete.

Proof of Statement 2): For $i = 1, 2, \cdots, N$, the discriminant Δ_i of Eq. (3.1.16)is defined as:

$$\Delta_i \triangleq -4\bar{\mathbf{x}}_i^T(a_{ii}\bar{\mathbf{A}}_{ii} - \bar{\mathbf{a}}_i\bar{\mathbf{a}}_i^T)\bar{\mathbf{x}}_i \tag{3.1.18}$$

Since it is either negative or equal to zero, we consider the following two cases.

1. $\Delta_i < 0$. In this case, $\bar{\mathbf{x}}_i \neq \mathbf{0}_{N\times 1}$. Otherwise, $\Delta_i = 0$. Then, $\forall \mathbf{x} \in \mathbb{R}_+^N$, we can have $a_{ii}x_i^2 + 2x_i\bar{\mathbf{a}}_i^T\bar{\mathbf{x}}_i + \bar{\mathbf{x}}_i^T\bar{\mathbf{A}}_{ii}\bar{\mathbf{x}}_i = 0$ with respect to x_i has no solution and thus, $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \neq 0$.

2. $\Delta_i = 0$. In this situation, $\bar{\mathbf{x}}_i^T(a_{ii}\bar{\mathbf{A}}_{ii} - \bar{\mathbf{a}}_i\bar{\mathbf{a}}_i^T)\bar{\mathbf{x}}_i = 0$. Since $(a_{ii}\bar{\mathbf{A}}_{ii} - \bar{\mathbf{a}}_i\bar{\mathbf{a}}_i^T)$ is full-cover, Theorem 3.1.1 tells that $\bar{\mathbf{x}}_i^T(a_{ii}\bar{\mathbf{A}}_{ii} - \bar{\mathbf{a}}_i\bar{\mathbf{a}}_i^T)\bar{\mathbf{x}}_i = 0$ if and only if $\bar{\mathbf{x}}_i = \mathbf{0}_{N-1}$.

Then,

$$\begin{aligned}\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} &= a_{ii}x_i^2 + 2x_i\bar{\mathbf{a}}_i^T\bar{\mathbf{x}}_i + \bar{\mathbf{x}}_i^T\bar{\mathbf{A}}_{ii}\bar{\mathbf{x}}_i \\ &= a_{ii}x_i^2\end{aligned} \tag{3.1.19}$$

Therefore, $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = 0$ if and only if $x_i = 0$. Combining this with $\bar{\mathbf{x}}_i = \mathbf{0}_{N-1}$ gives us that $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = 0$ if and only if $\mathbf{x} = \mathbf{0}_{N\times 1}$. By Theorem 3.1.1, we conclude that $\mathbf{A}^T\mathbf{A}$ is of full-cover.

This completes the sufficiency proof of Statement 2)

Proof of Statement 3): We prove this by contradiction. Suppose that $\mathbf{A}^T\mathbf{A}$ is of full-cover and there exists a principal sub-matrix $\hat{\mathbf{A}}$ of $\mathbf{A}^T\mathbf{A}$ such that $\hat{\mathbf{A}}$ is not of full-cover. Without loss of generality, we assume that $\hat{\mathbf{A}}$ is formed by the entries $a_{i_k^{(-)}i_l^{(-)}}$, where $k, \ell = 1, \cdots, n$ with $1 \leq n \leq N$. When $n = N$, we arrive at a contradiction to our assumption that $\mathbf{A}^T\mathbf{A}$ is of full-cover. When $1 \leq n < N$, by Theorem 3.1.1, there exists an $n \times n$ nonzero vector $\hat{\mathbf{x}}$ with nonnegative entries such that $\hat{\mathbf{x}}^T\hat{\mathbf{A}}\hat{\mathbf{x}} = 0$. Then, we construct an $N \times 1$ vector $\mathbf{x}$ by letting $x_{i_k^{(-)}} = \hat{x}_{i_k^{(-)}}$, for $k = 1, \cdots, n$ and the other $(N-n)$ entries of $\mathbf{x}$ being zeros. We arrive at:

$$\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = \hat{\mathbf{x}}^T\hat{\mathbf{A}}\hat{\mathbf{x}} = 0 \tag{3.1.20}$$

contradicting with our assumption that $\mathbf{A}^T\mathbf{A}$ is of full-cover. Therefore, Statement 3) is indeed necessary for $\mathbf{A}^T\mathbf{A}$ being of full-cover.

Proof of Statement 4): Again, we prove it by contradiction. We suppose that $\mathbf{A}^T\mathbf{A}$ is of full-cover and that there exists i such that $\mathbf{a}_i \notin \mathbf{R}_+^{N-1}$ and $\bar{\mathbf{C}}_i^{(-)}$ is not of full-cover. Then, by Theorem 3.1.1, there exists an $n \times 1$ non-zero vector $\bar{\mathbf{p}}$ with positive entries such that

$$\bar{\mathbf{p}}^T \bar{\mathbf{C}}_i^{(-)} \bar{\mathbf{p}} = 0 \tag{3.1.21}$$

Now, we form an $(N-1) \times 1$ vector $\bar{\mathbf{x}}_i$ by letting $\bar{x}_{i_1^{(-)}} = \bar{p}_1, \ \cdots \ , \bar{x}_{i_n^{(-)}} = \bar{p}_n$ and the other $(N-n-1)$ entries be zeros. In this case, we obtain

$$\begin{aligned} \Delta_i &= -4\bar{\mathbf{x}}_i^T (a_{ii}\bar{\mathbf{A}}_{ii} - \bar{\mathbf{a}}_i \bar{\mathbf{a}}_i^T)\bar{\mathbf{x}}_i \\ &= -4a_{ii}\bar{\mathbf{x}}_i^T \bar{\mathbf{A}}_{ii}\bar{\mathbf{x}}_i + 4\left(\bar{\mathbf{a}}_i^T \bar{\mathbf{x}}_i\right)^2 \\ &= -4a_{ii}\bar{\mathbf{p}}^T \bar{\mathbf{C}}_i^{(-)} \bar{\mathbf{p}} + 4\left(\sum_{i=i_1}^{i_{(N-1)}} \bar{a}_i \bar{x}_i\right)^2 \\ &= 0 + 4\left(\sum_{i=i_1^{(-)}}^{i_n^{(-)}} \bar{a}_i \bar{p}_i\right)^2 \\ &\geq 0 \end{aligned} \tag{3.1.22}$$

Using our notations, we know that $\bar{a}_{i_1^{(-)}} < 0, \ \cdots, \ \bar{a}_{i_n^{(-)}} < 0$ and thus

$$\bar{\mathbf{a}}_i^T \bar{\mathbf{x}}_i = \sum_{i=i_1^{(-)}}^{i_n^{(-)}} \bar{a}_i \bar{p}_i < 0 \tag{3.1.23}$$

Since $a_{ii} > 0$, $-\frac{\bar{\mathbf{a}}_i^T \bar{\mathbf{x}}_i}{a_{ii}} > 0$ and $\Delta_i > 0$, we can conclude that the quadratic equation $a_{ii}x_i^2 + 2x_i\bar{\mathbf{a}}_i^T \bar{\mathbf{x}}_i + \bar{\mathbf{x}}_i^T \bar{\mathbf{A}}_{ii}\bar{\mathbf{x}}_i = 0$ with respect to x_i has one positive solution given by

$$x_i = \frac{-\sum_{i=i_1^{(-)}}^{i_n^{(-)}} \bar{a}_i \bar{p}_i + \sqrt{\Delta_i}}{2a_{ii}} \tag{3.1.24}$$

Now, we see that there exists an $N \times 1$ non-zero vector $\mathbf{x} \in \mathbb{R}_+^N$ formed by x_i and $\bar{\mathbf{x}}_i$ such that $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = 0$. By Theorem 3.1.1, we know that $\mathbf{A}^T\mathbf{A}$ is not of full-cover, contradicting our assumption that $\mathbf{A}^T\mathbf{A}$ is of full-cover. Therefore, this completes the proof of Statement 4) as well as Theorem 3.1.2. □

In addition, for the block-diagonal PSD matrix, we have the following result of full cover.

Theorem 3.1.3. *Let $\mathbf{A}_\ell$ be an $N_\ell \times N_\ell$ PSD matrix for $\ell = 1, \cdots, L$. Then, the block diagonal PSD matrix $\mathbf{A} = diag(\mathbf{A}_1, \cdots, \mathbf{A}_L)$ has full cover if and only if $\mathbf{A}_\ell$ is of full-cover for any ℓ.* ■

Proof. Let $\mathbf{x}_\ell$ be a non-negative column vector with N_ℓ real elements in it, for $\ell = 1, \cdots, L$ and let

$$\mathbf{x} = \left(\mathbf{x}_1^T, \cdots, \mathbf{x}_L^T\right)^T \tag{3.1.25}$$

Then,

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{\ell=1}^{L} \mathbf{x}_\ell^T\mathbf{A}_l\mathbf{x}_l \tag{3.1.26}$$

On the one hand, if $\mathbf{A}_\ell$ is of full-cover for any ℓ, then, Theorem 3.2.2 tells us that $\mathbf{x}_\ell^T\mathbf{A}_\ell\mathbf{x}_\ell = 0$ if and only if $\mathbf{x}_\ell = \mathbf{0}_{N_\ell\times 1}$. This implies that $\mathbf{x}^T\mathbf{A}\mathbf{x} = 0$ if and only if $\mathbf{x} = \mathbf{0}$, thus by Theorem 3.1.1 we can obtain that $\mathbf{A}$ has full cover.

On the other hand, if $\mathbf{A}$ has full cover, then, according to Theorem 3.1.1, $\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{\ell=1}^{L} \mathbf{x}_\ell^T\mathbf{A}_\ell\mathbf{x}_\ell = 0$ if and only if $\mathbf{x}_\ell = \mathbf{0}$ for any ℓ. In addition, $\mathbf{x}_\ell^T\mathbf{A}_\ell\mathbf{x}_\ell \geq 0$ implies that $\sum_{\ell=1}^{L} \mathbf{x}_\ell^T\mathbf{A}_\ell\mathbf{x}_\ell = 0$ and $\mathbf{x}_\ell^T\mathbf{A}_\ell\mathbf{x}_\ell = 0$ are equivalent. Combining the two statements leads to the conclusion that $\mathbf{A}_\ell$ is of full-cover for any ℓ.

This completes the proof of Theorem 3.1.3. $\square$

3.2 System of Non-Homogeneous Linear Equations with Nonnegativity Constraints on Solutions

Consider the system of linear equations $\mathbf{Ax} = \mathbf{b}$ with nonnegativity constraint $\mathbf{x} \geq \mathbf{0}$, where $\mathbf{A}$ is an $M \times N$ real matrix and $\mathbf{b}$ is a $M \times 1$ real vector. This type of problem is frequently encountered in the field of signal and image processing, multispectral data handling, and more (20; 22; 14). There are mature methods for performing the analysis of the solution of the linear equations system without the requirement of nonnegativity constraints on solutions.

However, the existing methods for analyzing the existence and uniqueness of non-negative solutions to the system of linear equations are mainly concerned with discovering other associated problems rather than addressing the problem directly or the analysis of matrices with certain structures rather than the general case (58; 10; 32).

The classical result for analyzing the existence of non-negative solution to the linear equations system are based on the Farkas' Lemma (52), which is the underlying principle for linear programming duality and has been critical to the development of mathematical optimization. There are a number of slightly different (but equivalent) formulations of the Farkas' lemma in the literature. Here we present the most typical one among them (23).

Lemma 3.2.1 (Farkas' lemma)**.** *Let* $\mathbf{A} \in \mathbb{R}^{M\times N}$ *and* $\mathbf{b} \in \mathbb{R}^{M}$*. Then exactly one of the following two statements is true:*

1. *There exists an* $\mathbf{x} \in \mathbb{R}^{N}$ *such that* $\mathbf{x} \geq \mathbf{0}$ *and* $\mathbf{A}\mathbf{x} = \mathbf{b}$*.*

2. *There exists a* $\mathbf{y} \in \mathbb{R}^{M}$ *such that* $\mathbf{A}^{T}\mathbf{y} \geq \mathbf{0}$ *and* $\mathbf{b}^{T}\mathbf{y} < 0$

■

According to this lemma, given a problem of linear equations system with non-negativity constraints on the solutions, there exist another dual problem which is associated with it, and the original problem has a solution in the required domain if and only if the associated dual problem has no solution in the specific domain. To be more specific, Farkas' lemma shows that if a set of inequalities has no solution, then a contradiction can be produced from it by linear combination with non-negative coefficients. In formulas: if the inequalities in the option 2 cannot be satisfied, i.e., $\mathbf{A}^{T}\mathbf{y} \geq \mathbf{0}$, $\mathbf{b}^{T}\mathbf{y} < 0$ has no solution, then, the equality in option 1, $\mathbf{A}\mathbf{x} = \mathbf{b}$ has non-negative solutions. As for the conditions that guarantee the uniqueness of the non-negative solution, there is no direct characterization in general.

In the following, we will provide an investigation of the existence and uniqueness of non-negative solutions of the aforementioned linear equations system based on the cover theory. The necessary and sufficient conditions for the existence and the uniqueness of the non-negative solutions for such a system are derived.

Theorem 3.2.1. *Let* $\mathbf{A}$ *be an* $M \times N$ *real matrix. Then, a system of non-homogeneous linear equations with nonnegativity constraints on solutions:* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *has a solution in* $\mathbb{R}_{+}^{N}$ *if and only if the cover order of the matrix* $\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{A} & -\mathbf{b} \end{pmatrix}$ *is less than or equal to that of* $\mathbf{A}$*.* ■

Proof. By moving the right hand side constant vector $\mathbf{b}$ to the left hand side of the equation, the linear equation $\mathbf{Ax} = \mathbf{b}$ can be rewritten as $\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \mathbf{0}$, where the new coefficient matrix $\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{A} & -\mathbf{b} \end{pmatrix}$ and the new variable $\tilde{\mathbf{x}} = \begin{pmatrix} \mathbf{x} & \tilde{x}_{N+1} \end{pmatrix}$, and $\tilde{x}_{N+1}$ is the $(N+1)$-th element of $\tilde{\mathbf{x}}$, which equals to 1 exactly.

Let us first prove the sufficient condition. In this case, under the assumption $R_c(\tilde{\mathbf{A}}) \leq R_c(\mathbf{A})$, then by Statement 2) of Corollary 3.1.1, we can claim that $\tilde{x}_{N+1}$ is not covered in $\tilde{\mathbf{A}}\tilde{\mathbf{x}}$. Since $\tilde{\mathbf{A}}$ is not of full-cover, by Theorem 3.1.1, there exists a $\tilde{\mathbf{x}}_0 = \begin{pmatrix} \mathbf{x}_0 & \tilde{x}_{0,N+1} \end{pmatrix} \in \mathbb{R}_+^{N+1}$ 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

$$\mathbf{Ax}_0 = \mathbf{b}\tilde{x}_{0,N+1}, \tag{3.2.1}$$

Eq. (3.2.1) implies that $\mathbf{x}_0/\tilde{x}_{0,N+1}$ is a solution of $\mathbf{Ax} = \mathbf{b}$. Therefore, the sufficient condition is true.

Now, let us consider the proof of the necessary condition. In this case, we assume that a system of linear equations with nonnegativity constraints on solutions: $\mathbf{Ax} = \mathbf{b}$ has a solution in $\mathbb{R}_+^N$, i.e., there exists a $\mathbf{x}_0 \in \mathbb{R}_+^N$ such that the equality $\mathbf{Ax}_0 = \mathbf{b}$ holds. Then, according to Theorem 3.1.1, the $(N+1)$-th column vector of $\tilde{\mathbf{A}}$ is not covered. In addition, by Statement 3) of Corollary 3.1.1, we know that if the i-th column vector of $\tilde{\mathbf{A}}$, where $i \in \{1, 2, \cdots, N\}$, is covered in $\tilde{\mathbf{A}}$, then, as 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 the sufficient condition and thus, also completes the proof of Theorem 3.2.1. $\square$

The following example may give a clearer understanding of the necessary and sufficient condition for the existence of the non-negative solutions to the linear equations system.

Example 3.2.1. *Consider a linear equations system* $\mathbf{Ax} = \mathbf{b}$*, where*

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 3 \\ 4 \end{pmatrix},$$

and we have $R_c(\mathbf{A}) = 2$ *according to the definition of cover order in Definition 2.1.1. Then by adding the column vector* $-\mathbf{b}$ *to the right hand side of matrix* $\mathbf{A}$*, we will have the new coefficient matrix*

$$\tilde{\mathbf{A}} = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 1 & -4 \end{pmatrix}$$

The cover order of $\tilde{\mathbf{A}}$ *is 0, which is less than the cover order of* $\mathbf{A}$*, i.e.,* $R_c(\tilde{\mathbf{A}}) \leq R_c(\mathbf{A})$*. As a result, according to Theorem 3.2.1,* $\mathbf{Ax} = \mathbf{b}$ *has a solution in* $\mathbb{R}_+^N$ *and the solution is* $\mathbf{x} = \begin{pmatrix} \frac{5}{3} & \frac{2}{3} \end{pmatrix}^T$*.* ■

From Theorem 3.2.1, we can easily obtain the status of solutions to the non-homogeneous linear equations system with nonnegativity constraints on solutions under two specific scenarios. The results are stated in the following corollary.

Corollary 3.2.1. *Let* $\mathbf{A}$ *be an* $M \times N$ *real matrix and* $\mathbf{b}$ *be an* $M \times 1$ *real vector. Then, for a system of non-homogeneous linear equations* $\mathbf{Ax} = \mathbf{b}$*, with nonnegativity constraints on solutions* $\mathbf{x} \geq \mathbf{0}$*,*

1. *if the matrix* $\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{A} & -\mathbf{b} \end{pmatrix}$ *is full cover, then the system has no solution in* $\mathbb{R}_+^N$*;*

2. *if the matrix* $\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{A} & -\mathbf{b} \end{pmatrix}$ *is zero cover, then the system has at least one solution in* $\mathbb{R}_+^N$*.*

■

Furthermore, based on Theorem 3.2.1 and the above example, we can easily obtain the following important properties:

Property 3.2.1. *Given an* $M \times N$ *real matrix* $\mathbf{A}$*,*

1. *adding a column to it may decrease its cover order;*

2. *adding a row to it will not decrease its cover order.*

■

It is noted that Theorem 3.2.1 only specifies the necessary and sufficient condition for the existence of non-negative solutions for the system of non-homogeneous linear equations. With regard to the uniqueness of the solution for such a system, no direct analytical necessary and sufficient condition is applicable. This brings us to the next discussion on the uniqueness condition of the solution for such a linear equations system.

Our result is obtained by applying the following Carathéordory's theorem (52), which states that if a point $\mathbf{y}$ of $\mathbb{R}^M$ lies in the convex hull of a set $\mathcal{S}$, then $\mathbf{y}$ can be expressed as the convex combination of at most $M+1$ points in $\mathcal{S}$. Namely, there is a subset $\mathcal{T}$ of $\mathcal{S}$ consisting of $M+1$ or fewer points such that $\mathbf{y}$ lies in the convex hull of $\mathcal{T}$. The related definition of cone is given as follows (27).

Definition 3.2.1. *A set $\mathcal{C} \subseteq \mathbb{R}^M$ is a cone if $\alpha\mathbf{x} + \beta\mathbf{y} \in \mathcal{C}$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{C}$ and $\alpha, \beta > 0$.* ■

The following example provides us with a kind of cone spanned by a series of vectors:

Example 3.2.2. *Let $\mathbf{a}_1, \cdots, \mathbf{a}_N \in \mathbb{R}^M$. Then the cone spanned by the vectors $\mathbf{a}_1, \cdots, \mathbf{a}_N$ is*

$$\text{cone}\,\{\mathbf{a}_1, \cdots, \mathbf{a}_N\} = \{\theta_1\mathbf{a}_1 + \cdots + \theta_N\mathbf{a}_N | \theta_i \geq 0 \text{ for } i = 1, \cdots, N\}. \quad (3.2.2)$$

■

The Carathéordory's theorem that we are going to apply is given in the following lemma.

Lemma 3.2.2 (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 $\mathcal{T} \subseteq \mathcal{S}$ such that $\mathbf{y} \in \text{cone}\ \mathcal{T}$.* ■

We next show the necessary and sufficient condition for the uniqueness of the solution to the system of non-homogeneous linear equations which has nonnegativity constraints on solutions, with the aid of aforementioned Carathéordory's theorem. We state the theorem as follows.

Theorem 3.2.2. *Let $\mathbf{A}$ be an $M \times N$ real matrix. Then, a system of linear equations with nonnegativity constraints on solutions: $\mathbf{Ax} = \mathbf{b}$ has a unique solution in $\mathbb{R}_+^N$ if and only if $R_c(\tilde{\mathbf{A}}) \leq R_c(\mathbf{A})$ and $R_c(\tilde{\mathbf{A}}) + R_r(\bar{\mathbf{A}}) = N$, $\bar{\mathbf{A}} = \{\mathbf{a}_i\}_{i \in \bar{\mathcal{N}}}$, where $\bar{\mathcal{N}}$ is a set consisting of all the column indexes of $\mathbf{A}$ which are not covered in $\tilde{\mathbf{A}}$ and $R_r(\bar{\mathbf{A}})$ is the rank of $\bar{\mathbf{A}}$.* ■

Proof. As a first step, let us prove the sufficient condition. Let $\bar{R}_c(\tilde{\mathbf{A}})$ be the number of uncovered column vectors in $\tilde{\mathbf{A}}$. Thus we have

$$R_c(\tilde{\mathbf{A}}) + \bar{R}_c(\tilde{\mathbf{A}}) = N + 1 \tag{3.2.3}$$

and under the assumption $R_c(\tilde{\mathbf{A}}) + R_r(\bar{\mathbf{A}}) = N$, we have

$$\bar{R}_c(\tilde{\mathbf{A}}) = R_r(\bar{\mathbf{A}}) + 1 \tag{3.2.4}$$

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, a system of linear equations with nonnegativity constraints on solutions: $\mathbf{A}\mathbf{x} = \mathbf{b}$ has a unique solution.

To prove the necessary condition, we note that according to the property of $R_c(\tilde{\mathbf{A}})$ and the definition of $\bar{\mathbf{A}}$, we have

$$R_c(\tilde{\mathbf{A}}) + R_r(\bar{\mathbf{A}}) \leq N \tag{3.2.5}$$

Let us assume $\mathbf{x}_0 \in \mathbb{R}_+^N$ is the unique solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$, then we have $\mathbf{A}\mathbf{x}_0 = \mathbf{b}$, where

$$x_{0,i} = 0, \text{ for } i \in \{1, 2, \cdots, N\} \setminus \bar{\mathcal{N}} \tag{3.2.6}$$

Considering $R_c(\tilde{\mathbf{A}}) + R_r(\bar{\mathbf{A}}) < N$, in this case, the rank of $\bar{\mathbf{A}}$ must be less than $|\bar{\mathcal{N}}|$, i.e., the inequality $R_r(\bar{\mathbf{A}}) < |\bar{\mathcal{N}}|$ must be satisfied. As the 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 exists another solution $\mathbf{x}_1 \in \mathbb{R}_+^N$ and $\mathbf{x}_1 \neq \mathbf{x}_0$ such that $\mathbf{A}\mathbf{x}_1 = \mathbf{b}$ and we have

$$x_{1,i} = 0, \text{ for } i \in \{1, 2, \cdots, N\} \setminus \tilde{\mathcal{N}} \tag{3.2.7}$$

This statement contradicts with the assumption that $\mathbf{x}_0$ is the unique solution of the linear equation $\mathbf{A}\mathbf{x} = \mathbf{b}$. As a result, the equality $R_c(\tilde{\mathbf{A}}) + R_r(\bar{\mathbf{A}}) = N$ must be satisfied to guarantee the uniqueness of the solution to the linear equation system with nonnegativity constraints on variables. Thus, Theorem 3.2.2 is proved. □

Example 3.2.3. *Consider linear equations system* $\mathbf{A}\mathbf{x} = \mathbf{b}$*, where*

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

and $R_c(\mathbf{A}) = 2$*. Then we will have*

$$\tilde{\mathbf{A}} = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \end{pmatrix}$$

In order to determine whether a unique solution exists for such a system, we first need to confirm that a solution exists. Since $R_c(\tilde{\mathbf{A}}) = 0$*, which is less than the cover order of the origin coefficient matrix* $\mathbf{A}$*. According to Theorem 3.2.1,* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *has a solution in* $\mathbb{R}_+^N$*. In the next step, we need to determine whether it meets the uniqueness requirement in Theorem 3.2.2. In this example,* $\bar{\mathbf{A}}$ *is equivalent to* $\mathbf{A}$ *exactly, thus* $R_r(\bar{\mathbf{A}}) = R_r(\mathbf{A}) = 1$*. Then we have* $R_c(\tilde{\mathbf{A}}) + R_r(\bar{\mathbf{A}}) = 1$*, which is less than the* *column number of* $\mathbf{A}$*. As a consequence, the solution of this linear equations system is not unique. Moreover, it is easy to obtain two explicit solutions:* $\begin{pmatrix} 1 & 1 \end{pmatrix}^T$ *and* $\begin{pmatrix} 0 & \frac{3}{2} \end{pmatrix}^T$. ■

Specifically, in Example 3.2.1, we note that $R_r(\bar{\mathbf{A}}) = R_r(\mathbf{A}) = 2$, and thus we have $R_c(\tilde{\mathbf{A}}) + R_r(\bar{\mathbf{A}}) = 2$. As a result, the uniqueness condition in Theorem 3.2.2 is satisfied and the linear equations system with nonnegativity constraint has a unique solution, which is $\mathbf{x} = \begin{pmatrix} \frac{5}{3} & \frac{2}{3} \end{pmatrix}^T$.

3.3 Conclusion

Given an $M \times N$ real matrix $\mathbf{A}$ and the real column vector $\mathbf{x}$ with N entries in it, we propose a method of determining whether the i-th column of $\mathbf{A}$ or the i-th variable x_i corresponding to the i-th column vector $\mathbf{a}_i$ in $\mathbf{Ax}$ is covered or not. In the process of this verification, we are able to establish the equivalence relationship between full cover and the unique solution of homogeneous linear equations system which has nonnegativity constraints on solutions. We also introduce some related algebraic conditions on full cover in order to have a better understanding of the properties of cover order. Additionally, we also derive the necessary and sufficient condition that guarantees the existence of non-zero solutions to the system of linear equations with nonnegativity constraints. By using cover theory, we also obtain the necessary and sufficient condition that leads to the unique solution for the system of linear equations with nonnegativity constraints. It is noteworthy that these conditions are based on the matrix itself, meaning that we can analyze the solution of the system of linear equations with non-negative constraints on solutions from the matrix itself rather than from other associated problems.

In comparison with the linear equations system without the nonnegativity constraints on variables, we can summarize the results as follows.

- For homogeneous linear equations:

  1. $\mathbf{Ax} = \mathbf{0}$ has a nonzero solution if and only if rank($\mathbf{A}$)$< N$
  2. $\mathbf{Ax} = \mathbf{0}$ has a nonzero solution in $\mathbb{R}_+^N$ if and only if $R_c(\mathbf{A}) < N$
  3. $\mathbf{Ax} = \mathbf{0}$ has a unique solution if and only if rank($\mathbf{A}$)$= N$
  4. $\mathbf{Ax} = \mathbf{0}$ has a unique solution in $\mathbb{R}_+^N$ if and only if $R_c(\mathbf{A}) = N$

- For non-homogeneous linear equations:

  1. $\mathbf{Ax} = \mathbf{b}$ has a solution if and only if rank($\mathbf{A}, \mathbf{b}$)=rank($\mathbf{A}$)
  2. $\mathbf{Ax} = \mathbf{b}$ has a solution in $\mathbb{R}_+^N$ if and only if $R_c(\mathbf{A}, -\mathbf{b}) \leq R_c(\mathbf{A})$
  3. $\mathbf{Ax} = \mathbf{b}$ has a unique solution if and only if rank($\mathbf{A}, \mathbf{b}$)=rank($\mathbf{A}$)$= N$
  4. $\mathbf{Ax} = \mathbf{b}$ has a unique solution in $\mathbb{R}_+^N$ if and only if $R_c(\mathbf{A}, -\mathbf{b}) \leq R_c(\mathbf{A})$ and $R_c(\mathbf{A}, -\mathbf{b}) + R_r(\bar{\mathbf{A}}) = N$, where $\bar{\mathbf{A}}$ is defined in Theorem 3.2.2

Chapter 4

Cover Order

In this chapter, we present the necessary and sufficient condition to determine the cover order of any given real matrix. With this condition, we are motivated to establish a specific échelon form of the matrix by applying a series of elementary row operations and column permutations to it. This échelon form enables us to determine the cover order for any given matrix efficiently and effectively. For zero-cover matrices, we also reveal the special structure and discover some properties of them. Furthermore, we investigate the specific échelon form for some special matrices. In the process of investigating the échelon form of the matrix, we arrive at a more profound understanding in the analysis of linear equations system with nonnegativity constraints on solutions.

Based on the échelon form of matrix, the concepts of non-negatively linear independence and non-negatively linear dependence have been developed. From these concepts, we acquire a deeper insight into the system of linear equations with non-negativity constraints on solutions. Then, we introduce the related concepts on the construction of the generalized inverse of a matrix and based on Farkas' lemma, we propose a new dual property on the coverage of a matrix.

Moreover, in the same behavior that we defined the hyper-rectangle that covers the set $\left\{\mathbf{x} : 0 \leq \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2, \mathbf{x} \in \mathbb{R}_+^N\right\}$ in Chapter 2, we introduce the inner hyper-rectangle in the set $\left\{\mathbf{x} : 0 \leq \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2, \mathbf{x} \in \mathbb{R}_+^N\right\}$, which is shown to have a strong relationship with the properties of zero-cover.

4.1 Properties of Cover Order

To ensure the completeness of the discussion, we present here some results related to orthogonal complementary extracted from (1; 51), which can be applied to derive the necessary and sufficient condition to determine the cover order of any given real matrix. In the following discussion, $\mathbb{R}_{++}^N$ denotes the set of all the $N \times 1$ vectors with all N entries being positive and $\mathbb{R}_+^N$ denotes the set of all the $N \times 1$ vectors with all N entries being non-negative. If all the elements in the vector $\mathbf{x}$ are positive , then we say, it is a positive vector. If all the elements in the vector $\mathbf{x}$ are all non-negative, then it is a non-negative vector. Similarily, we can define the negative vector and the non-positive vector correspondingly.

The definition of orthogonal complement and the related properties are given in the following.

Definition 4.1.1 (Orthogonal Complement)**.** *The orthogonal complement of a subspace $\mathbb{S}$ of a vector space $\mathbb{V}$ equipped with a bilinear form B is the set $\mathbb{S}^\perp$ of all vectors in $\mathbb{V}$ that are orthogonal to every vector in $\mathbb{S}$.* ■

Lemma 4.1.1. *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* $\mathbb{S}^{\perp} \cap \mathbb{R}_{+}^{N} \neq \emptyset$.

■

The definition of the row space of any matrix is given as follows (2; 53).

Definition 4.1.2 (Row Space)**.** *Let* $\mathcal{K}$ *be a field of scalars. Let* $\mathbf{A}$ *be an* $M \times N$ *matrix, with row vectors* $\mathbf{r}_1,\ \mathbf{r}_2,\ \cdots,\mathbf{r}_M$. *A linear combination of these vectors is any vector of the form*

$$c_1\mathbf{r}_1 + c_2\mathbf{r}_2 + \cdots + c_M\mathbf{r}_M,$$

where $c_1,\ c_2,\ \cdots,\ c_M$ *are scalars. The set of all possible linear combinations of* $\mathbf{r}_1,\ \mathbf{r}_2,\ \cdots,\mathbf{r}_M$ *is called the row space of* $\mathbf{A}$. *That is, the row space of* $\mathbf{A}$ *is the span of the vectors* $\mathbf{r}_1,\ \mathbf{r}_2,\ \cdots,\mathbf{r}_M$. ■

Denote the row space of $\mathbf{A}$ by $\mathbb{S}_{\mathbf{A}}$ and the orthogonal complement to this row space by $\mathbb{S}_{\mathbf{A}}^{\perp}$. Using Lemma 4.1.1, the following necessary and sufficient condition to determine the cover order of a matrix is developed, which is stated as the following theorem.

Theorem 4.1.1. *Let* $\bar{\mathbb{R}}_{++}^{K}$ *denote the set of all the nonnegative vectors with* K *positive entries and specifically,* $\bar{\mathbb{R}}_{++}^{0}$ *be* $\{\mathbf{0}_{N\times 1}\}$. *For any* $\mathbf{A} \in \mathbb{R}^{M\times N}$, *the cover order of* $\mathbf{A}$ *is equal to* $\max_{\mathbb{S}_{\mathbf{A}} \cap \bar{\mathbb{R}}_{++}^{K} \neq \emptyset} K$. ■

Proof. The proof is done by first verifying $R_c \geq \max_{\mathbb{S}_{\mathbf{A}} \cap \bar{\mathbb{R}}_{++}^{K}}$ and then proving $R_c \leq \max_{\mathbb{S}_{\mathbf{A}} \cap \bar{\mathbb{R}}_{++}^{K}}$.

Proof of $R_c \geq \max_{\mathbb{S}_\mathbf{A} \cap \bar{\mathbb{R}}^K_{++}}$: Let us denote the i-th row of $\mathbf{A}$ by $\mathbf{a}_i^T$ for $i = 1, 2, \cdots, M$ and first assume

$$K_0 = \max_{\mathbb{S}_\mathbf{A} \cap \bar{\mathbb{R}}^K_{++} \neq \emptyset} K \geq 1 \tag{4.1.1}$$

By this assumption, there exists an $M \times 1$ vector $\mathbf{v}$ such that

$$\mathbf{v}^T \mathbf{A} = \mathbf{p}^T \in \mathbb{S}_\mathbf{A} \cap \bar{\mathbb{R}}^{K_0}_{++} \tag{4.1.2}$$

For an arbitrarily given $\tau > 0$, from $\mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} = \|\mathbf{A}\mathbf{x}\|_2^2 \leq \tau^2$ we have, for $i = 1, \cdots, M$: $-\tau \leq \mathbf{a}_i^T \mathbf{x} \leq \tau$, where $\mathbf{x} \in \mathbb{R}^N_+$. Notice that $\forall v_i \in \mathbb{R}, i = 1, 2, \cdots, M$, we have

$$-|v_i|\tau \leq v_i \mathbf{a}_i^T \mathbf{x} \leq |v_i|\tau \tag{4.1.3}$$

Summing the above M inequalities, we can obtain

$$-\tau \sum_{i=1}^{M} |v_i| \leq \mathbf{p}^T \mathbf{x} \leq \tau \sum_{i=1}^{M} |v_i| \tag{4.1.4}$$

Since $\mathbf{p}^T \in \mathbb{S}_\mathbf{A} \cap \bar{\mathbb{R}}^{K_0}_{++}$, there exist K_0 integers $\ell_k^{(+)} \in \{1, \ldots, N\}$ for $k = 1, \cdots, K_0$ such that $p_{\ell_k}^{(+)} > 0$ by our definition of $\bar{\mathbb{R}}^{K_0}_{++}$ and thus, we have $\mathbf{p}^T \mathbf{x} \geq 0$ for $\mathbf{x} \geq \mathbf{0}$. Combining this observation with (4.1.4) yields

$$0 \leq x_{\ell_k^{(+)}} \leq \frac{\tau}{p_{\ell_k^{(+)}}} \sum_{i=1}^{M} |v_i|, \qquad k = 1, \cdots, K_0 \tag{4.1.5}$$

Thus,

$$\begin{aligned} &\{\mathbf{x} : \mathbf{x} \in \mathbb{R}_+^N, \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \le \tau^2\} \\ &\subseteq \left\{\mathbf{x} : 0 \le x_{\ell_k^{(+)}} \le \frac{\sum_{i=1}^M |v_i|}{p_{\ell_k^{(+)}}}\tau,\ k = 1,\ \cdots,\ K_0\right\} \end{aligned} \tag{4.1.6}$$

Then, according to the definition of cover order in Definition 2.1.1, we have

$$R_c \ge K_0 \tag{4.1.7}$$

When $K_0 = N$, we may have $R_c = N$ since the definition of cover order in Definition 2.1.1 tells us that $0 \le R_c \le N$.

Proof of $R_c \le \max_{\mathbb{S}_{\mathbf{A}} \cap \bar{\mathbb{R}}_{++}^K}$: In the following, we consider the case with $R_c < N$. Let $\mathbf{p}$ be defined in (4.1.2), which is an $N \times 1$ nonnegative vector with K_0 positive entries. Denoting the indices of the zero-valued entries of $\mathbf{p}$ by $\ell_k^{(0)}$, where $k = 1,\ \cdots,\ (N - K_0)$, we obtain

$$\{\ell_k^{(0)}, k = 1,\ \cdots,\ (N - K_0)\} \cap \{\ell_k^{(+)}, k = 1,\ \cdots\ K_0\} = \emptyset \tag{4.1.8}$$

We can now generate an $M \times (N - K_0)$ sub-matrix of $\mathbf{A}$ by using the $(N - K_0)$ columns of $\mathbf{A}$ indexed by $\ell_k^{(0)}, k = 1,\ \cdots,\ (N - K_0)$ and denote this sub-matrix by $\bar{\mathbf{A}}^{(0)}$. We claim that $\mathbb{S}_{\bar{\mathbf{A}}^{(0)}} \cap \mathbb{R}_+^{N-K_0} = \emptyset$. This claim can be proved by contradiction: Let us first suppose that

$$\mathbb{S}_{\bar{\mathbf{A}}^{(0)}} \cap \mathbb{R}_+^{N-K_0} \ne \emptyset \tag{4.1.9}$$

As a consequence, there exists an $(N-K_0)\times 1$ vector $\bar{\mathbf{q}}^{(0)}$ such that

$$\bar{\mathbf{q}}^{(0)} \in \mathbb{S}_{\bar{\mathbf{A}}^{(0)}} \cap \mathbb{R}_{+}^{N-K_0} \tag{4.1.10}$$

Then, according to the definition of $\bar{\mathbf{A}}^{(0)}$, there exists a vector $\tilde{\mathbf{p}}$ such that the $\ell_k^{(0)}$-th entry of $\tilde{\mathbf{p}}$ is given by the k-th element of $\bar{\mathbf{q}}^{(0)}$ for all $k=1, \cdots, (N-K_0)$ and $\tilde{\mathbf{p}} \in \mathbb{S}_{\mathbf{A}}$. Then, we denote the minimum of the K_0 positive entries of $\mathbf{p}$ (defined in (4.1.2)) indexed by $\ell_k^{(+)}, k=1, \cdots, K_0$ by $p_{\min}^{(+)}$. Moreover, let us denote the minimum and the maximum of the entries of $\tilde{\mathbf{p}}$ indexed by $\ell_k^{(+)}, k=1, \cdots, K_0$ by $\tilde{p}_{\min}^{(+)}$ and $\tilde{p}_{\max}^{(+)}$, respectively. We notice that for $\ell_k^{(+)}$ with $k=1, \cdots, K_0$,

$$\begin{aligned}&\frac{|\tilde{p}_{\min}^{(+)}|+|\tilde{p}_{\max}^{(+)}|+1}{p_{\min}^{(+)}} p_{\ell_k^{(+)}} + \tilde{p}_{\ell_k^{(+)}} \\ &\geq |\tilde{p}_{\min}^{(+)}|+|\tilde{p}_{\max}^{(+)}|+1+\tilde{p}_{\ell_k^{(+)}} \geq 1 > 0\end{aligned} \tag{4.1.11}$$

by the definitions of $p_{\min}^{(+)}$, $\tilde{p}_{\min}^{(+)}$ and $\tilde{p}_{\max}^{(+)}$. From $\bar{\mathbf{q}}^{(0)} \in \mathbb{S}_{\bar{\mathbf{A}}^{(0)}} \cap \mathbb{R}_{+}^{N-K_0}$, where $\mathbb{R}_{+}^{N-K_0}$ is defined by the $(N-K_0)\times 1$ nonnegative vector with at least one positive entry, we can have that among all the $(N-K_0)$ nonnegative entries of $\tilde{\mathbf{p}}$ indexed by $\ell_k^{(0)}$ for $k=1, \cdots, (N-K_0)$, at least one is positive. Combining this observation with the above established facts that

$$\{\ell_k^{(0)}, 1 \leq k \leq (N-K_0)\} \cap \{\ell_k^{(+)}, 1 \leq k \leq K_0\} = \emptyset \tag{4.1.12}$$

and the entries of $\frac{|\tilde{p}_{\min}^{(+)}|+|\tilde{p}_{\max}^{(+)}|+1}{p_{\min}^{(+)}}\mathbf{p}+\tilde{\mathbf{p}}$ indexed by $\ell_k^{(+)}, k=1, \cdots, K_0$ are positive, we can conclude that at least (K_0+1) entries of $\frac{|\tilde{p}_{\min}^{(+)}|+|\tilde{p}_{\max}^{(+)}|+1}{p_{\min}^{(+)}}\mathbf{p}+\tilde{\mathbf{p}}$ are positive. Putting this result and $\mathbf{p},\ \tilde{\mathbf{p}} \in \mathbb{S}_{\mathbf{A}}$ together leads us to

$$\frac{|\tilde{p}_{\min}^{(+)}| + |\tilde{p}_{\max}^{(+)}| + 1}{p_{\min}^{(+)}}\mathbf{p} + \tilde{\mathbf{p}} \in \mathbb{S}_{\mathbf{A}} \cap \bar{\mathbb{R}}_{++}^{K} \tag{4.1.13}$$

with $K \geq K_0 + 1$, which contradicts to the assumption that $K_0 = \max_{\mathbb{S}_{\mathbf{A}} \cap \bar{\mathbb{R}}_{++}^{K} \neq \emptyset} K$ in (4.1.1). Thus, our assumption that $\mathbb{S}_{\bar{\mathbf{A}}^{(0)}} \cap \mathbb{R}_{+}^{N-K_0} \neq \emptyset$ in (4.1.9) is not true, and we arrive at

$$\mathbb{S}_{\bar{\mathbf{A}}^{(0)}} \cap \mathbb{R}_{+}^{N-K_0} = \emptyset \tag{4.1.14}$$

In addition, by Lemma 4.1.1, if $\mathbb{S}_{\bar{\mathbf{A}}^{(0)}} \cap \mathbb{R}_{+}^{N-K_0} = \emptyset$, then, there exists an $(N - K_0) \times 1$ vector $\tilde{\mathbf{q}}^{(0)}$ such that

$$\tilde{\mathbf{q}}^{(0)} \in \mathbb{S}_{\bar{\mathbf{A}}^{(0)}}^{\perp} \cap \mathbb{R}_{++}^{N-K_0} \tag{4.1.15}$$

Then, we generate an $N \times 1$ vector $\mathbf{x}_0$ by letting $x_{\ell_k^{(+)}} = 0$ for $k = 1, \ \cdots, \ K_0$ and $x_{\ell_k^{(0)}} = \tilde{q}_k^{(0)}$ for $k = 1, \ \cdots, \ (N - K_0)$, respectively. Since $\tilde{\mathbf{q}}^{(0)} \in \mathbb{S}_{\bar{\mathbf{A}}^{(0)}}^{\perp} \cap \mathbb{R}_{++}^{N-K_0}$ gives us $\bar{\mathbf{A}}^{(0)}\tilde{\mathbf{q}}^{(0)} = \mathbf{0}$, we obtain

$$\mathbf{x}_0^T \mathbf{A}^T \mathbf{A} \mathbf{x}_0 = 0 + (\tilde{\mathbf{q}}^{(0)})^T (\bar{\mathbf{A}}^{(0)})^T \bar{\mathbf{A}}^{(0)} \tilde{\mathbf{q}}^{(0)} = 0 \tag{4.1.16}$$

Now, for any positive τ, the following relationship holds

$$\{\xi \mathbf{x}_0 : \xi > 0\} \subset \{\mathbf{x} : \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} \leq \tau^2, \mathbf{x} \in \mathbb{R}_{+}^{N}\} \tag{4.1.17}$$

Since the positive quantity ξ is arbitrary, then for any $\tau > 0$, there exists no positive constant $c_{\ell_k^{(0)}}$ such that

$$\xi x_{\ell_k^{(0)}} \leq c_{\ell_k^{(0)}} \tau \tag{4.1.18}$$

In other words, at least $(N - K_0)$ entries of $\mathbf{x}_0$ can not be covered. By the definition of cover order in Definition 2.1.1, which says that the cover order is the maximum number of the covered entries of any given nonzero $\mathbf{x}$, we arrive at $R_c \leq K_0$.

Combining $R_c \leq K_0$ with (4.1.7) yields $R_c = K_0$. For the case $K_0 = 0$, following the same argument for $R_c \leq K_0$, we can have $R_c \leq 0$. By Definition 2.1.1, R_c is a nonnegative integer and thus, we can conclude that $R_c = K_0 = 0$. The proof of Theorem 4.1.1 is thus complete. □

According to Theorem 4.1.1, we can easily obtain the following corollary and lemma.

Corollary 4.1.1. *Given an $M \times N$ real matrix $\mathbf{A}$, if there exists at least one non-zero row vector in $\mathbf{A}$ in which all the elements have the same sign, then, $\mathbf{A}$ has a full-cover.* ■

Lemma 4.1.2. *Given an $M \times M$ real square matrix $\mathbf{A}$, if, for every column of the matrix, there exists an absolute value of an entry in a column larger than the sum of all the absolute value of other entries in the same column, and they lie in different rows, i.e., if*

$$|a_{ij}| > \sum_{k \neq i} |a_{kj}| \qquad \text{for all } j$$

where a_{ij}, $i \in \{1, 2, \cdots, M\}$, denotes the entry in the i-th row and j-th column in *the matrix* $\mathbf{A}$*, then* $\mathbf{A}$ *has a full-cover.* ■

Proof. To simplify the expression, we named the entry in a column whose absolute value is larger than the sum of the absolute value of other entries in this column as the column dominant entry. Suppose for the j-th column, the column dominant entry is a_{ij}, where $j = 1,\ 2,\ \cdots,\ M$ and $i \in \{1, 2, \cdots, M\}$. Denote the row vectors of $\mathbf{A}$ as $\mathbf{r}_1,\ \mathbf{r}_2,\ \cdots,\ \mathbf{r}_M$. Then we can find a linear combination of these row vector of the form:

$$\mathbf{r} = c_1\mathbf{r}_1 + c_2\mathbf{r}_2 + \cdots + c_M\mathbf{r}_M,$$

where $c_j = 1$ if the column dominant entry of the j-th column a_{ij} is positive and $c_j = -1$ if $a_{ij} < 0$. In this case, $\mathbf{r}$ is a positive row vector with M elements. Thus, according to Theorem 4.1.1, we have, $\mathbf{A}$ has a full-cover. □

We also derive an important property of cover order which is stated in detail as the following theorem.

Theorem 4.1.2. *Given an* $M \times N$ *real matrix* $\mathbf{A}$*, doing row operations to* $\mathbf{A}$ *will not change its cover order. Specifically, given an invertible* $M \times M$ *real matrix* $\mathbf{T}$*, let* $\mathbf{B} = \mathbf{TA}$*, 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 $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum eigenvalue and the maximum eigenvalue of $\mathbf{T}^T\mathbf{T}$ respectively, then we have

$$\lambda_{\min}\mathbf{I} \leq \mathbf{T}^T\mathbf{T} \leq \lambda_{\max}\mathbf{I} \tag{4.1.19}$$

According to above inequalities, we have

$$\lambda_{\min}(\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x}) \leq \mathbf{x}^T(\mathbf{A}^T\mathbf{T}^T\mathbf{T}\mathbf{A})\mathbf{x} \leq \lambda_{\max}(\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x}) \tag{4.1.20}$$

Then, using the left hand side inequality: $\lambda_{\min}(\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x}) \leq \mathbf{x}^T\mathbf{B}^T\mathbf{B}\mathbf{x}$ in Eq. (4.1.20) and the definition of cover order in Definition 2.1.1, we have, for any real-valued number $\tau > 0$,

$$\begin{aligned} \{\mathbf{x} : \mathbf{x} \in \mathbb{R}_+^N, \mathbf{x}^T\mathbf{B}^T\mathbf{B}\mathbf{x} \leq \tau^2\} &\subseteq \{\mathbf{x} : \mathbf{x} \in \mathbb{R}_+^N, \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \frac{\tau^2}{\lambda_{min}}\} \\ &\subseteq \{\mathbf{x} : 0 \leq x_{k_i} \leq \frac{c_{k_i}}{\lambda_{min}}\tau, i = 1, \cdots, R_c(\mathbf{A})\} \end{aligned}$$

where all c_{k_i} are absolutely constants independent of τ, then by the definition of cover order in Definition 2.1.1, we know that at least $R_c(\mathbf{A})$ variables in $\mathbf{x}$ are covered. Thus we have $R_c(\mathbf{A}) \leq R_c(\mathbf{B})$.

According to the right hand side inequality in Eq. (4.1.20), i.e., $\mathbf{x}^T\mathbf{B}^T\mathbf{B}\mathbf{x} \leq \lambda_{\max}(\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x})$, we have, for any real-valued number $\tau > 0$,

$$\begin{aligned} \{\mathbf{x} : \mathbf{x} \in \mathbb{R}_+^N, \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \frac{\tau^2}{\lambda_{max}}\} &\subseteq \{\mathbf{x} : \mathbf{x} \in \mathbb{R}_+^N, \mathbf{x}^T\mathbf{B}^T\mathbf{B}\mathbf{x} \leq \tau^2\} \\ &\subseteq \{\mathbf{x} : 0 \leq x_{k_i} \leq c_{k_i}\tau, i = 1, \cdots, R_c(\mathbf{B})\} \end{aligned}$$

where all c_{k_i} are constants independent of τ, then with the definition of cover order in Definition 2.1.1, we know that at least $R_c(\mathbf{B})$ variables in $\mathbf{x}$ are covered. Thus we have $R_c(\mathbf{A}) \geq R_c(\mathbf{B})$.

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

According to Theorem 4.1.1, we know that if we are able to find nonnegative 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. This result together with Theorem 4.1.2 indeed implicitly suggests that we can perform a series of linear elementary row transformations and column permutations to determine the cover order of the matrix $\mathbf{A}$. This suggestion leads us to the following section in which we introduce an échelon transformation of the matrix to obtain the cover order of any given matrix.

4.2 Échelon Form of Matrix

4.2.1 Échelon Transformation

Given an $M \times N$ real matrix $\mathbf{A}$, the process of determining the cover order of $\mathbf{A}$ can be described in the following steps:

1) *Échelon form.* Given an $M \times N$ real matrix $\mathbf{A}$, we can find the elementary transformation matrix $\mathbf{E_0}$ and permutation matrix $\mathbf{P_0}$ such that

$$\mathbf{E_0 A P_0} = \begin{pmatrix} \mathbf{I} & \mathbf{B_0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{4.2.1}$$

where $\mathbf{I} \in \mathbb{R}^{R_r \times R_r}$, $\mathbf{B_0} \in \mathbb{R}^{R_r \times (N-R_r)}$ and R_r is the rank of the matrix $\mathbf{A}$. Eq. (4.2.1) is called the *échelon form* of $\mathbf{A}$. It is noted that the transformation towards the matrix $\mathbf{A}$ is not unique, we can choose different $\mathbf{E}_0$ and $\mathbf{P}_0$ arriving at different values of $\mathbf{B}_0$.

Without loss of generality, we can assume $\mathbf{A}$ is of full-rank, i.e., $R_r = M$. In particular, from Theorem 3.1.1 and Theorem 4.1.1,

(a) if the initial échelon form transformation of $\mathbf{A}$ in Eq. (4.2.1) results in every entry in some *row* of $\mathbf{B}_0$ being positive, then, $R_c(\mathbf{A}) = N$, i.e, $\mathbf{A}$ has a full-cover;

(b) 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 form, the following steps of structural arrangement can be taken to determine the cover order.

2) *Structure Arrangement.* Search for all non-negative rows in $\mathbf{B_0}$ and select the one which has the greatest number of positive elements. Moving this selected row to the first row and assuming that it contains N_1 positive entries. By doing the row and column permutation, we can always ensure the identity matrix structure ahead and let the following statements hold:

$$b_{11}, b_{12}, \cdots, b_{1N_1} > 0,\ b_{1(N_1+1)}, \cdots, b_{1(N-M)} = 0 \tag{4.2.2}$$

where b_{1i}, $i = 1, 2, \ \cdots, \ N - M$ are the elements in the first row of the new structure of $\mathbf{B_0}$. Then, in the above, ignoring the first N_1 columns in new $\mathbf{B_0}$. 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 it contains N_2 positive entries in the remaining $N - M - N_1$ columns, we have:

$$b_{2(N_1+1)}, b_{2(N_1+2)}, \cdots, b_{2(N_1+N_2)} > 0, b_{2(N_1+N_2+1)}, \cdots, b_{2(N-M)} = 0 \tag{4.2.3}$$

where b_{2i}, $i = N_1+1, N_1+2, \cdots, N-M$, are the elements in the second row of the new form of $\mathbf{B_0}$ after above steps. By arranging the following rows similarly, after s times, we obtain:

$$\begin{aligned}
b_{11}, b_{12}, \cdots, b_{1N_1} &> 0, \\
b_{1(N_1+1)}, \cdots, b_{1(N-M)} &= 0; \\
b_{2(N_1+1)}, b_{2(N_1+2)}, \cdots, b_{2(N_1+N_2)} &> 0, \\
b_{2(N_1+N_2+1)}, \cdots, b_{2(N-M)} &= 0; \\
&\vdots \\
b_{(s-1)(N_1+N_2+ \cdots +N_{s-2}+1)}, \cdots, b_{(s-1)(N_1+ \cdots +N_(s-1)} &> 0, \\
b_{(s-1)(N_1+N_2+ \cdots +N_{(s-1)}+1)}, \cdots, b_{(s-1)(N-M)} &= 0; \\
b_{s(N_1+N_2+ \cdots +N_{s-1}+1)}, \cdots, b_{s(N_1+ \cdots +N_s)} &> 0, \\
b_{s(N_1+N_2+ \cdots +N_s+1)}, \cdots, b_{s(N-M)} &= 0.
\end{aligned} \tag{4.2.4}$$

where b_{ij} in Eqs. (4.2.4), $i = 1, 2, \cdots, s$, $j = 1, 2, \cdots, N-M$, and $s \leq M$, are the elements in the first s rows of the structure of $\mathbf{B_0}$ after s times transformation. Then let:

$$\bar{\mathbf{B}} = \begin{pmatrix} b_{s+1,N_1+\cdots+N_s+1} & b_{s+1,N_1+\cdots+N_s+2} & \cdots & b_{s+1,N-M} \\ b_{s+2,N_1+\cdots+N_s+1} & b_{s+2,N_1+\cdots+N_s+2} & \cdots & b_{s+2,N-M} \\ \vdots & \vdots & \ddots & \vdots \\ b_{M-1,N_1+\cdots+N_s+1} & b_{M-1,N_1+\cdots+N_s+2} & \cdots & b_{M-1,N-M} \\ b_{M,N_1+\cdots+N_s+1} & b_{M,N_1+\cdots+N_s+2} & \cdots & b_{M,N-M} \end{pmatrix} \tag{4.2.5}$$

where b_{ij} in Eq. (4.2.5), $i = s+1, s+2, \cdots, M$, $j = N_1+\cdots+N_s+1, N_1+$

$\cdots + N_s + 2, \ \cdots, \ N - M$, and $s \leq M$, are the elements in the bottom-right corner of the structure of $\mathbf{B_0}$ after s times transformation.

The row and column transformations on $\mathbf{A}$ ends when one of the following two cases happens:

(a) $\sum_{i=1}^{s} N_i = N - M$, in which case, $\mathbf{A}$ is full cover;

(b) There is no non-negative row vector in the row space of $\bar{\mathbf{B}}$.

3) *Cover Order*. At the end of the above structural arrangement procedure, we arrive at the conclusion that the cover order of $\mathbf{A}$ is

$$R_c(\mathbf{A}) = \sum_{i=1}^{s} N_i + s \tag{4.2.6}$$

where $s \leq M$.

■

The next theorem states the property of the final form of the échelon form of any given matrix.

Theorem 4.2.1. *For any $M \times N$ real matrix $\mathbf{A}$, there exists elementary matrix $\mathbf{E}$ and permutation matrix $\mathbf{P}$, such that:*

$$\mathbf{EAP} = \begin{pmatrix} \mathbf{I} & \mathbf{B} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{4.2.7}$$

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; or*

2. *contains at least one negative column vector, or there exists one non-positive column vector, but the position where the zero lies will be negative in some other columns of* **B**.

■

More specifically, when the first scenario occurs, i.e., **B** has at least one non-negative row in it, then the cover order of **A** can be determined by the above steps. Otherwise, $R_c(\mathbf{A}) = 0$.

In the following, we present the proof of Theorem 4.2.1.

Proof. As stated before, without lose of generality, we can assume **A** has full rank, i.e., $R_r = M$.

1. For $N - M = 1$, i.e.,

$$\mathbf{EAP} = \begin{pmatrix} \mathbf{I} & \mathbf{b} \end{pmatrix}, \tag{4.2.8}$$

   where **b** is a $M \times 1$ column vector. 1) If **b** contains at least one positive elements, then **A** is of full-cover. 2) If **b** is negative vector, then **A** is of zero-cover. 3) If **b** is a non-positive vector, then the cover order of **A** equals to the number of zero term.

2. Suppose for $N - M = K$, the above conclusion holds: i.e., if $R_c(\mathbf{A}) > 0$, then

$\mathbf{A}$ can be transformed into the following form:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1K} \\ 0 & 1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2K} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & b_{(M-1)1} & b_{(M-1)2} & \cdots & b_{(M-1)K} \\ 0 & 0 & \cdots & 1 & b_{M1} & b_{M2} & \cdots & b_{MK} \end{pmatrix} \tag{4.2.9}$$

where

$$\begin{aligned} b_{11}, b_{12}, \cdots, b_{1N_1} > 0, \\ b_{1(N_1+1)}, \cdots, b_{1K} = 0; \\ b_{2(N_1+1)}, b_{2(N_1+2)}, \cdots, b_{2(N_1+N_2)} > 0, \\ b_{2(N_1+N_2+1)}, \cdots, b_{2K} = 0; \\ \vdots \\ b_{(s-1)(N_1+N_2+\ \cdots\ +N_{s-2}+1)},\ \ \cdots\ , b_{(s-1)(N_1+\ \cdots\ +N_(s-1)} > 0, \\ b_{(s-1)(N_1+N_2+\ \cdots\ +N_{(s-1)}+1)},\ \ \cdots\ , b_{(s-1)K} = 0; \\ b_{s(N_1+N_2+\cdots+N_{s-1}+1)}, \cdots, b_{s(N_1+\cdots+N_s)} > 0, \\ b_{s(N_1+N_2+\cdots+N_s+1)}, \cdots, b_{sK} = 0 \end{aligned} \tag{4.2.10}$$

and $R_c(\mathbf{A}) = \sum_{i=1}^{s} N_i + s$. Besides, let

$$\bar{\mathbf{B}} = \begin{pmatrix} b_{s+1,N_1+\cdots+N_s+1} & \cdots & b_{s+1,K} \\ \vdots & \ddots & \vdots \\ b_{M,N_1+\cdots+N_s+1} & \cdots & b_{M,K} \end{pmatrix} \tag{4.2.11}$$

where $\bar{\mathbf{B}}$ either contains at least one negative column vector, or has one non-positive column vector, 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 $N - M = K + 1$.

3. When $N - M = K + 1$, assuming

$$\mathbf{A} = \begin{pmatrix} \mathbf{I} & \mathbf{b}_1 & \cdots & \mathbf{b}_K & \mathbf{b}_{K+1} \end{pmatrix} \tag{4.2.12}$$

where $\mathbf{I}$ is a $M \times M$ identity matrix and $\mathbf{b}_i, i = 1, \cdots, K+1$ are all $M \times 1$ column vectors. Without considering $\mathbf{b}_{K+1}$, denote the remaining part in $\mathbf{A}$ as $\bar{\mathbf{A}}$, which equals to

$$\bar{\mathbf{A}} = \begin{pmatrix} \mathbf{I} & \mathbf{b}_1 & \cdots & \mathbf{b}_K \end{pmatrix} \tag{4.2.13}$$

According to the assumption in the case of $N - M = K$, $\bar{\mathbf{A}}$ can be transformed into échelon form and let

$$R_c(\bar{\mathbf{A}}) = \sum_{i=1}^{s} N_i + s \tag{4.2.14}$$

By considering the corresponding $\mathbf{b}_{K+1}$ with the éhelon 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 see that if $b_{1,K+1} > 0$, and we have:

$$R_c(\mathbf{A}) = \sum_{i=1}^{s} N_i + s + 1 \tag{4.2.15}$$

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 $b_{1,K+1} < 0$. The following steps can be taken to make the first row non-positive and move it to the last row without affecting the cover order of $\mathbf{A}$.

Step 1 Let:

$$m = \max_{j \in \{1,2,\ \cdots,\ M\}} \{-\frac{b_{j1}}{b_{11}}, -\frac{b_{j2}}{b_{12}}\ \cdots,\ -\frac{b_{j,K+1}}{b_{1,K+1}}\}, \tag{4.2.16}$$

where $b_{j1}b_{11} < 0$, $b_{j2}b_{12} < 0,\ \cdots,\ b_{j,K+1}b_{1,K+1} < 0$, for $j = 1, 2,\ \cdots,\ M$. Using m times the first row of $\mathbf{A}$ and adding the product to each row in it, then we have $\mathbf{A}_{(1)}$:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & b_{11} & \cdots & b_{1,K+1} \\ m & 1 & \cdots & 0 & b_{21}+mb_{11} & \cdots & b_{2,K+1}+mb_{1,K+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ m & 0 & \cdots & 1 & b_{M1}+mb_{11} & \cdots & b_{M,K+1}+mb_{1,K+1} \end{pmatrix} \tag{4.2.17}$$

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

$$\begin{pmatrix} -1 & 0 & \cdots & 0 & -b_{11} & \cdots & -b_{1,K+1} \\ m & 1 & \cdots & 0 & b_{21}+mb_{11} & \cdots & b_{2,K+1}+mb_{1,K+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ m & 0 & \cdots & 1 & b_{M1}+mb_{11} & \cdots & b_{M,K+1}+mb_{1,K+1} \end{pmatrix} \tag{4.2.18}$$

Step 3 Let:

$$\begin{aligned} t_2 &= \frac{b_{2,K+1}}{b_{1,K+1}} + m, \\ t_3 &= \frac{b_{3,K+1}}{b_{1,K+1}} + m, \\ &\vdots \\ t_M &= \frac{b_{M,K+1}}{b_{1,K+1}} + m. \end{aligned} \tag{4.2.19}$$

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 have $\mathbf{A}_{(3)}$:

$$\begin{pmatrix} -1 & \cdots & 0 & -b_{11} & \cdots & -b_{1,K+1} \\ m - t_2 & \cdots & 0 & \frac{\begin{vmatrix} b_{21} & b_{2,K+1} \\ b_{11} & b_{1,K+1} \end{vmatrix}}{b_{1,K+1}} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ m - t_M & \cdots & 1 & \frac{\begin{vmatrix} b_{M1} & b_{M,K+1} \\ b_{11} & b_{1,K+1} \end{vmatrix}}{b_{1,K+1}} & \cdots & 0 \end{pmatrix} \tag{4.2.20}$$

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

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & \frac{b_{11}}{b_{1,K+1}} & \cdots & \frac{1}{b_{1,K+1}} \\ 0 & 1 & \cdots & 0 & \frac{\begin{vmatrix} b_{21} & b_{2,K+1} \\ b_{11} & b_{1,K+1} \end{vmatrix}}{b_{1,K+1}} & \cdots & -\frac{b_{2,K+1}}{b_{1,K+1}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \frac{\begin{vmatrix} b_{M1} & b_{M,K+1} \\ b_{11} & b_{1,K+1} \end{vmatrix}}{b_{1,K+1}} & \cdots & -\frac{b_{M,K+1}}{b_{1,K+1}} \end{pmatrix} \tag{4.2.21}$$

Step 5 Permuting the corresponding rows and columns such 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, we have $\mathbf{A}_{(5)}$.

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & \frac{\begin{vmatrix} b_{21} & b_{2,K+1} \\ b_{11} & b_{1,K+1} \end{vmatrix}}{b_{1,K+1}} & \cdots & -\frac{b_{2,K+1}}{b_{1,K+1}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \frac{\begin{vmatrix} b_{M1} & b_{M,K+1} \\ b_{11} & b_{1,K+1} \end{vmatrix}}{b_{1,K+1}} & \cdots & -\frac{b_{M,K+1}}{b_{1,K+1}} \\ 0 & 0 & \cdots & 1 & \frac{b_{11}}{b_{1,K+1}} & \cdots & \frac{1}{b_{1,K+1}} \end{pmatrix} \tag{4.2.22}$$

Step 6 Without considering the last column of $\mathbf{A}_{(5)}$, we rearrange the rows and columns of the first $(M+K)$ columns such that we are able to obtain a new échelon form matrix $\bar{\mathbf{A}}_{(5)}$ and

$$R_c(\bar{\mathbf{A}}_{(5)}) = \sum_{i=1}^{s^{(2)}} N_i^{(2)} + s^{(2)} \tag{4.2.23}$$

By considering the corresponding $\bar{\mathbf{b}}_{K+1}$ in the échelon form of $\bar{\mathbf{A}}_{(5)}$, we see that if $\bar{b}_{1,K+1} > 0$, then

$$R_c(\mathbf{A}) = \sum_{i=1}^{s^{(2)}} N_i^{(2)} + s^{(2)} + 1 \tag{4.2.24}$$

If $\bar{b}_{1,K+1} < 0$, we can repeat the above steps.

Finally, either after t times transformation, there exists one $\bar{b}_{1,K+1} > 0$, such that the first row of the new matrix is non-negative and

$$R_c(\mathbf{A}) = \sum_{i=1}^{s^{(t)}} N_i^{(t)} + s^{(t)} + 1 \tag{4.2.25}$$

or, $R_c(\mathbf{A}) = 0$ and there exists at least one column ($(K+1)$-th column of $\mathbf{A}$) is negative.

□

To demonstrate how to determine the cover order of a given matrix based on the échelon form, the following examples are given.

Example 4.2.1. *The matrix* $\mathbf{A}$ *in the following examples is already in its échelon form. And based on this form, we can determine the cover order of* $\mathbf{A}$ *directly.*

1. *Consider the matrix:*

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 & -1 \end{pmatrix}$$

From the above discussion, by considering the first non-negative row in $\mathbf{A}$*, we can obtain that* $\mathbf{a}_1, \mathbf{a}_3$ *and* $\mathbf{a}_4$ *are all covered column vectors in* $\mathbf{A}$*. Then without looking at those covered columns, there is no non-negative row in the remaining part of* $\mathbf{A}$*. Thus* $\mathbf{a}_2$*,* $\mathbf{a}_5$ *are uncovered. Correspondingly, the associated* x_1, x_3 *and* x_4 *in* $\mathbf{Ax}$ *are covered variables. As a result,* $R_c(\mathbf{A}) = 3$*.*

2. Consider the matrix:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 & -1 \end{pmatrix}$$

We note that there is a non-positive column vector $\mathbf{a}_5$ *in matrix* $\mathbf{A}$*. Besides, the row where the zero lies in* $\mathbf{a}_5$ *has negative element in the other column. As a consequence,* $\mathbf{A}$ *is zero cover.*

■

With the proof of Theorem 4.2.1, we can easily reach the conclusion that the zero-cover matrix possesses the following property.

Corollary 4.2.1. *For any* $M \times N$ *real matrix* $\mathbf{A}$*, if* $\mathbf{A}$ *is zero cover then there exists elementary matrix* $\mathbf{E}$ *and permutation matrix* $\mathbf{P}$*, such that:*

$$\mathbf{EAP} = \begin{pmatrix} \mathbf{I} & \mathbf{B} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{4.2.26}$$

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* $\mathbf{A}$*. The matrix*

B *possesses the following special structure:*

$$
\begin{aligned}
b_{1,1}, b_{2,1}, \cdots, b_{r_1,1} &< 0; \\
b_{r_1+1,1}, b_{r_1+2,1}, \cdots, b_{R_r,1} &= 0; \\
b_{r_1+1,2}, b_{r_1+2,2}, \cdots, b_{r_1+r_2,2} &< 0; \\
b_{r_1+r_2+1,2}, b_{r_1+r_2+2,2} \cdots, b_{R_r,2} &= 0; \\
&\vdots \\
b_{r_1+r_2+\cdots+r_{k-1}+1,k}, \cdots, b_{r_1+r_2+\cdots+r_{k-1}+r_k,k} &< 0.
\end{aligned}
$$

where $b_{i,j}$ *is the* (i,j)*-th element of matrix* **B**, $1 \leq k \leq (N - R_r)$ *and* $r_1 + r_2 + \cdots + r_{k-1} + r_k = R_r$.

■

4.2.2 Specific Échelon Form for Special Cases

Observe that the échelon form of a matrix is not unique, and under different circumstances, a different form may be required. Due to this, it is necessary to investigate the specific échelon form for special cases.

Lemma 4.2.1. *Consider an* $M \times N$ *full rank real matrix* **A** *and* $N \geq 2M$*. Suppose there is only one negative element in each column and each row in the right-hand side of its échelon form, while the remaining elements are all positive. If* **A** *has zero-cover, then* **A** *can be transformed into the following form, where the right hand side matrix* **B** *in* **A** *will contain one row only has one negative term while other rows are all* *non-positive vectors:*

$$\mathbf{A} \to \begin{pmatrix} 1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1,N-M} \\ 0 & 1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2,N-M} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & b_{M1} & b_{M2} & \cdots & b_{M,N-M} \end{pmatrix} \tag{4.2.27}$$

where $b_{11} < 0$ *while* $b_{1j} > 0$, *for* $j = 2, \cdots, N-M$ *and the remaining elements are all non-positive.* ■

Proof. Without loss of generality, we can assume the matrix $\mathbf{A}$ has full rank, i.e., $R_r = M$, and let

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1,N-M} \\ 0 & 1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2,N-M} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & b_{M1} & b_{M2} & \cdots & b_{M,N-M} \end{pmatrix} \tag{4.2.28}$$

where $b_{ii} < 0$ and $b_{ij} > 0$, for $i = 1,\ 2,\ \cdots,\ M$ and $j = 1,\ 2,\ \cdots,\ N-M$ and $N \geq 2M$.

Step 1 Let:

$$t_1 = -\frac{b_{1M}}{b_{MM}}, t_2 = -\frac{b_{2M}}{b_{MM}}, \cdots, t_{M-1} = -\frac{b_{M-1,M}}{b_{MM}}.$$

According to the assumption, we have $t_i > 0$ for $i = 1\ , 2,\ \cdots,\ M-1$. Let $\mathbf{a}_j^T$ be the j-th row of $\mathbf{A}$. Then by adding $\mathbf{a}_M^T t_j$ to the j-the row in $\mathbf{A}$, where

$j = 1\ ,2,\ \cdots,\ M-1$. We will have $\mathbf{A}^{(1)}$:

$$\begin{pmatrix} 1 & \cdots & t_1 & b_{11}+t_1 b_{M1} & \cdots & 0 & \cdots & b_{1,N-M}+t_1 b_{M,N-M} \\ 0 & \cdots & t_2 & b_{21}+t_2 b_{M1} & \cdots & 0 & \cdots & b_{2,N-M}+t_2 b_{M,N-M} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & b_{M1} & \cdots & b_{MM} & \cdots & b_{M,N-M} \end{pmatrix}$$

Step 2 Multiplying the last row of $\mathbf{A}^{(1)}$ by (-1) to get $\mathbf{A}^{(2)}$:

$$\begin{pmatrix} 1 & \cdots & t_1 & b_{11}+t_1 b_{M1} & \cdots & 0 & \cdots & b_{1,N-M}+t_1 b_{M,N-M} \\ 0 & \cdots & t_2 & b_{21}+t_2 b_{M1} & \cdots & 0 & \cdots & b_{2,N-M}+t_2 b_{M,N-M} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & -1 & -b_{M1} & \cdots & -b_{MM} & \cdots & -b_{M,N-M} \end{pmatrix}$$

Step 3 Multiplying the last row of $\mathbf{A}^{(2)}$ that we obtained in the last step by $-\frac{1}{b_{MM}}$ and exchanging the position of the M-th column with the $2M$-th column in $\mathbf{A}^{(2)}$ at the mean time, we will get $\mathbf{A}^{(3)}$:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & b_{11}+t_1 b_{M1} & \cdots & t_1 & \cdots & b_{1,N-M}+t_1 b_{M,N-M} \\ 0 & 1 & \cdots & 0 & b_{21}+t_2 b_{M1} & \cdots & t_2 & \cdots & b_{2,N-M}+t_2 b_{M,N-M} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & & \\ 0 & 0 & \cdots & 1 & \frac{b_{M1}}{b_{MM}} & \cdots & \frac{1}{b_{MM}} & \cdots & \frac{b_{M,N-M}}{b_{MM}} \end{pmatrix}$$

It is observed that in $\mathbf{A}^{(3)}$, the elements in the last row of right hand side of it are all negative, i.e.,

$$\begin{aligned} \frac{1}{b_{MM}} &< 0, \\ \frac{b_{Mj}}{b_{MM}} &< 0 \text{ for j} = 1, \ \cdots, \ \text{M} - 1, \ \text{M} + 1, \ \cdots, \ \text{N} - \text{M} \end{aligned} \tag{4.2.29}$$

In addition, in the remaining part of the right hand side of the matrix $\mathbf{A}^{(3)}$ we have:

(a) for $i \neq j$ where $i = 1, \ \cdots, \ M-1$ and $j = 1, \ \cdots, \ M-1, \ M+1, \ \cdots, \ N-M$:

$$b_{ij} + t_i b_{Mi} > 0 \tag{4.2.30}$$

(b) for $i = j$ where $i = 1, \ \cdots, \ M-1$:

$$b_{ii} + t_i b_{Mi} < 0 \tag{4.2.31}$$

The inequality (4.2.31) holds since the matrix $\mathbf{A}$ has zero cover, then each row of $\mathbf{A}^{(3)}$ must contain at least one negative element.

Step 4 Denote the elements in the right hand side of $\mathbf{A}^{(3)}$ as $b_{ij}^{(1)}$ for $i = 1, \ \cdots, \ M-1$ and $j = 1, \ \cdots, \ N-M$. Then let:

$$t_1^{(1)} = -\frac{b_{1,M-1}^{(1)}}{b_{M-1,M-1}^{(1)}}, t_2^{(1)} = -\frac{b_{2,M-1}^{(1)}}{b_{M-1,M-1}^{(1)}}, \cdots, t_{M-2}^{(1)} = -\frac{b_{M-2,M-1}^{(1)}}{b_{M-1,M-1}^{(1)}}, t_M^{(1)} = -\frac{b_{M,M-1}^{(1)}}{b_{M-1,M-1}^{(1)}}$$

According to the above discussions, we have

$$\begin{aligned} t_i^{(1)} &> 0 \text{ for } i = 1\ ,2,\ \cdots,\ M-2, \\ t_M^{(1)} &< 0. \end{aligned} \tag{4.2.32}$$

Let $\mathbf{a}_j^{(1)T}$ be the j-th row of $\mathbf{A}^{(3)}$. Then by adding $\mathbf{a}_j^{(1)T} t_i^{(1)}$ to the j-the row in $\mathbf{A}^{(3)}$, where $j = 1\ ,2,\ \cdots,\ M-2,\ M$, we will have $\mathbf{A}_{(4)}$:

$$\begin{pmatrix} 1 & \cdots & t_1^{(1)} & 0 & b_{11}^{(1)} + t_1^{(1)} b_{M-1,1}^{(1)} & \cdots & 0 & \cdots & b_{1,N-M}^{(1)} + t_1^{(1)} b_{M-1,N-M}^{(1)} \\ 0 & \cdots & t_2^{(1)} & 0 & b_{21}^{(1)} + t_2^{(1)} b_{M-1,1}^{(1)} & \cdots & 0 & \cdots & b_{2,N-M}^{(1)} + t_2^{(1)} b_{M-1,N-M}^{(1)} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & b_{M-1,1}^{(1)} & \cdots & b_{M-1,M-1}^{(1)} & \cdots & b_{M-1,N-M}^{(1)} \\ 0 & \cdots & t_M^{(1)} & 1 & b_{M1}^{(1)} + t_M^{(1)} b_{M-1,1}^{(1)} & \cdots & 0 & \cdots & b_{M,N-M}^{(1)} + t_M^{(1)} b_{M-1,N-M}^{(1)} \end{pmatrix}$$

Step 5 Multiply the $(M-1)$-th row of $\mathbf{A}^{(4)}$ by (-1), then multiplying the $(M-1)$-th row of it with $-\frac{1}{b_{M-1,M-1}}$ and exchanging the position of the $(M-1)$-th column with the $2(M-1)$-th column, we will get $\mathbf{A}^{(5)}$:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & b_{11}^{(1)} + t_1^{(1)} b_{M-1,1}^{(1)} & \cdots & t_1^{(1)} & \cdots & b_{1,N-M}^{(1)} + t_1 b_{M-1,N-M}^{(1)} \\ 0 & 1 & \cdots & 0 & b_{21}^{(1)} + t_2^{(1)} b_{M-1,1}^{(1)} & \cdots & t_2^{(1)} & \cdots & b_{2,N-M}^{(1)} + t_2^{(1)} b_{M-1,N-M}^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & & \\ 0 & 0 & \cdots & 0 & \frac{b_{M-1,1}^{(1)}}{b_{M-1,M-1}^{(1)}} & \cdots & \frac{1}{b_{M-1,M-1}^{(1)}} & \cdots & \frac{b_{M-1,N-M}^{(1)}}{b_{M-1,M-1}^{(1)}} \\ 0 & 0 & \cdots & 1 & b_{M1}^{(1)} + t_M^{(1)} b_{M-1,1}^{(1)} & \cdots & t_M^{(1)} & \cdots & b_{M,N-M}^{(1)} + t_M^{(1)} b_{M-1,N-M}^{(1)} \end{pmatrix}$$

We notice that in $\mathbf{A}^{(5)}$, the elements in the $(M-1)$-th row of right hand side of it are all negative, i.e.,

$$\begin{aligned} \frac{1}{b^{(1)}_{M-1,M-1}} &< 0, \\ \frac{b^{(1)}_{M-1,j}}{b^{(1)}_{M-1,M-1}} &< 0 \text{ for j} = 1, \ \cdots, \ \mathrm{M}-2, \ \mathrm{M}, \ \cdots, \ \mathrm{N}-\mathrm{M} \end{aligned} \quad (4.2.33)$$

Moreover, the last row of the right hand side of the matrix $\mathbf{A}^{(5)}$ are all negative as well, i.e.,

$$\begin{aligned} t^{(1)}_M &< 0, \\ b^{(1)}_{Mj} + t^{(1)}_M b^{(1)}_{M-1,j} &< 0 \text{ for j} = 1, \ \cdots, \mathrm{M}-2, \mathrm{M}, \cdots, \mathrm{N}-\mathrm{M} \end{aligned} \quad (4.2.34)$$

Additionally, in the remaining part of the right hand side of the matrix $\mathbf{A}^{(5)}$ we have:

(a) for $i \neq j$ where $i = 1, \ \cdots, \ M-2$ and $j = 1, \ \cdots, \ M-2, \ M, \ \cdots, \ N-M$:

$$b_{ij} + t_i b_{Mi} > 0 \quad (4.2.35)$$

(b) for $i = j$ where $i = 1, \ \cdots, \ M-2$:

$$b_{ii} + t_i b_{Mi} < 0 \quad (4.2.36)$$

The above inequality holds since the matrix $\mathbf{A}$ has zero cover, then each row of $\mathbf{A}^{(5)}$ must contain at least one negative element.

Repeat the whole process until the second row of the right hand side part of the new matrix has been transformed into a negative vector, and the final form is:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & b_{11}^{(M-1)} & b_{12}^{(M-1)} & \cdots & b_{1,N-M}^{(M-1)} \\ 0 & 1 & \cdots & 0 & b_{21}^{(M-1)} & b_{22}^{(M-1)} & \cdots & b_{2,N-M}^{(M-1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & b_{M1}^{(M-1)} & b_{M2}^{(M-1)} & \cdots & b_{M,N-M}^{(M-1)} \end{pmatrix}$$

where

$$\begin{aligned} b_{11}^{(M-1)} &< 0, \\ b_{1j}^{(M-1)} &> 0, \ \text{for j} = 1,\ 2,\ \cdots,\ \text{N} - \text{M} \\ b_{ij}^{(M-1)} &< 0, \ \text{for i} = 2,\ 3,\ \cdots,\ \text{M},\ \text{j} = 1,\ 2,\ \cdots,\ \text{N} - \text{M} \end{aligned}$$

Therefore, the proof of Lemma 4.2.1 is complete. □

From the proof of Lemma 4.2.1, we observe that this zero-cover matrix **A** contains one negative column vector in it, i.e., the elements in the $(M+1)$-th column of the matrix **A** possesses the special structure in the statement of Lemma 4.2.1, which are all negative.

Lemma 4.2.2. *Given a real matrix* **A***, if the matrix* $\mathbf{B_0}$ *in the échelon form of it only has two columns, then* **A** *can be transformed into the following form:*

$$\mathbf{EAP} = \begin{pmatrix} \mathbf{I} & \mathbf{B} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{4.2.37}$$

where the row vector of **B** *is either non-positive or non-negative. Specifically, if* **A** *has zero-cover, then* **B** *will be a non-positive matrix, besides, there is no zero row vector in* **B**. ■

Proof. According to the Theorem 4.2.1, the right-hand side matrix **B** in the échelon form of **A** will either

1. have at least one non-negative row vector; or

2. have at least one negative column or has one non-positive column but the position where the zero lies will be negative in another column.

Now let us suppose **A** has full rank and there exist the elementary transformation matrix **E** and the permutation matrix **P** such that

$$\mathbf{EAP} = \begin{pmatrix} \mathbf{I} & \mathbf{B} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 & b_{11} & b_{12} \\ 0 & 1 & \cdots & 0 & b_{21} & b_{22} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & b_{M1} & b_{M2} \end{pmatrix} \tag{4.2.38}$$

For the first case, i.e., when **B** contains at least one non-negative row vector. Without loss of generality, we can suppose b_{11} and b_{12} are both positive, while the combination of elements in other row vectors will be a) both negative, b) one positive and one negative.

Step 1 By performing the row and column permutation, we are able to move the row of **B** whose elements are both negative to the lower side of it while guaranteeing the identity matrix structure ahead, then we will have $\mathbf{A}^{(1)}$:

$$\mathbf{A}^{(1)} = \begin{pmatrix} \mathbf{I} & \mathbf{B}^{(1)} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 & b_{11}^{(1)} & b_{12}^{(1)} \\ 0 & 1 & \cdots & 0 & b_{21}^{(1)} & b_{22}^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & b_{M1}^{(1)} & b_{M2}^{(1)} \end{pmatrix}$$

Step 2 Let the number of negative row vectors in $\mathbf{B}^{(1)}$ be P and let $s = M-P$. Suppose $b_{s1}^{(1)} > 0$ and $b_{s2}^{(1)} < 0$, then let:

$$m = \max_{j\in\{1,\ 2,\ \cdots,\ M-P\}} \{-\frac{b_{j1}^{(1)}}{b_{s1}^{(1)}}, -\frac{b_{j2}^{(1)}}{b_{s2}^{(1)}}\}, \tag{4.2.39}$$

where $b_{j1}^{(1)} b_{s1}^{(1)} < 0$ and $b_{j2}^{(1)} b_{s2}^{(1)} < 0$ for $j = 1,\ 2,\ \cdots,\ M-P$.

Multiply the $(M-P)$-th row of $\mathbf{A}^{(1)}$ by m and adding the product to each row in it. Then we will have $\mathbf{A}^{(2)}$:

$$\begin{pmatrix} 1 & 0 & \cdots & m & \cdots & 0 & b_{11}^{(1)} + mb_{s1}^{(1)} & b_{12}^{(1)} + mb_{s2}^{(1)} \\ 0 & 1 & \cdots & m & \cdots & 0 & b_{21}^{(1)} + mb_{s1}^{(1)} & b_{22}^{(1)} + mb_{s2}^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & b_{s1}^{(1)} & b_{s2}^{(1)} \\ 0 & 0 & \cdots & 0 & \cdots & 0 & b_{s+1,1}^{(1)} & b_{s+1,2}^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & b_{M1}^{(1)} & b_{M2}^{(1)} \end{pmatrix}$$

It is observed that

$$b_{i1}^{(1)} + mb_{s1}^{(1)} > 0,\ b_{i2}^{(1)} + mb_{s2}^{(1)} < 0 \text{ for i} = 1,\ \cdots,\ \text{s} - 1;$$

$$b_{s1}^{(1)} > 0,\ b_{s2}^{(1)} < 0;$$

$$b_{i1}^{(1)} < 0,\ b_{i2}^{(1)} < 0 \text{ for i} = \text{s} + 1,\ \cdots,\ \text{M}.$$

Step 3 Multiply the s-th row of $\mathbf{A}^{(2)}$ by (-1) to get $\mathbf{A}^{(3)}$:

$$\begin{pmatrix} 1 & 0 & \cdots & m & \cdots & 0 & b_{11}^{(1)} + mb_{s1}^{(1)} & b_{12}^{(1)} + mb_{s2}^{(1)} \\ 0 & 1 & \cdots & m & \cdots & 0 & b_{21}^{(1)} + mb_{s1}^{(1)} & b_{22}^{(1)} + mb_{s2}^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & \cdots & 0 & -b_{s1}^{(1)} & -b_{s2}^{(1)} \\ 0 & 0 & \cdots & 0 & \cdots & 0 & b_{s+1,1}^{(1)} & b_{s+1,2}^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & b_{M1}^{(1)} & b_{M2}^{(1)} \end{pmatrix}$$

Step 4 Then for $i = 1,\ 2,\ \cdots,\ s-1$:

$$t_i = \frac{b_{i2}^{(1)}}{b_{s2}^{(1)}} + m, \tag{4.2.40}$$

and for $i = s+1,\ s+2,\ \cdots,\ M$:

$$t_i = \frac{b_{i2}^{(1)}}{b_{s2}^{(1)}}. \tag{4.2.41}$$

Let $\mathbf{a}_j{}^T$ be the j-th row of $\mathbf{A}^{(3)}$. Then, by adding $\mathbf{a}_j{}^T t_j$ to the j-the row in $\mathbf{A}^{(3)}$, where $j = 1, \ \cdots, \ s-1, \ s+1 \ \cdots, \ M$, we have $\mathbf{A}^{(4)}$:

$$
\begin{pmatrix}
1 & 0 & \cdots & -\frac{b_{12}^{(1)}}{b_{s2}^{(1)}} & \cdots & 0 & \frac{\begin{vmatrix} b_{11}^{(1)} & b_{12}^{(1)} \\ b_{s1}^{(1)} & b_{s2}^{(1)} \end{vmatrix}}{b_{s2}^{(1)}} & 0 \\
0 & 1 & \cdots & -\frac{b_{22}^{(1)}}{b_{s2}^{(1)}} & \cdots & 0 & \frac{\begin{vmatrix} b_{21}^{(1)} & b_{22}^{(1)} \\ b_{s1}^{(1)} & b_{s2}^{(1)} \end{vmatrix}}{b_{s2}^{(1)}} & 0 \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & -1 & \cdots & 0 & -b_{s1}^{(1)} & -b_{s2}^{(1)} \\
0 & 0 & \cdots & -t_{s+1,2} & \cdots & 0 & \frac{\begin{vmatrix} b_{s+1,1}^{(1)} & b_{s+1,2}^{(1)} \\ b_{s1}^{(1)} & b_{s2}^{(1)} \end{vmatrix}}{b_{s2}^{(1)}} & 0 \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & -t_{M2} & \cdots & 1 & \frac{\begin{vmatrix} b_{M1}^{(1)} & b_{M2}^{(1)} \\ b_{s1}^{(1)} & b_{s2}^{(1)} \end{vmatrix}}{b_{s2}^{(1)}} & 0
\end{pmatrix}
$$

Step 5 Multiplying the s-th row of $\mathbf{A}^{(4)}$ obtained in the last step with $-\frac{1}{b_{s2}}$ and exchanging the position of the s-th column with the last column, we have $\mathbf{A}^{(5)}$:

$$\begin{pmatrix}
1 & 0 & \cdots & 0 & \cdots & 0 & \frac{\begin{vmatrix} b_{11}^{(1)} & b_{12}^{(1)} \\ b_{s1}^{(1)} & b_{s2}^{(1)} \end{vmatrix}}{b_{s2}^{(1)}} & -\frac{b_{12}^{(1)}}{b_{s2}^{(1)}} \\
0 & 1 & \cdots & 0 & \cdots & 0 & \frac{\begin{vmatrix} b_{21}^{(1)} & b_{22}^{(1)} \\ b_{s1}^{(1)} & b_{s2}^{(1)} \end{vmatrix}}{b_{s2}^{(1)}} & -\frac{b_{22}^{(1)}}{b_{s2}^{(1)}} \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & 1 & \cdots & 0 & \frac{b_{s1}^{(1)}}{b_{s2}^{(1)}} & \frac{1}{b_{s2}^{(1)}} \\
0 & 0 & \cdots & 0 & \cdots & 0 & \frac{\begin{vmatrix} b_{s+1,1}^{(1)} & b_{s+1,2}^{(1)} \\ b_{s1}^{(1)} & b_{s2}^{(1)} \end{vmatrix}}{b_{s2}^{(1)}} & -\frac{b_{s+1,2}^{(1)}}{b_{s2}^{(1)}} \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & \cdots & 1 & \frac{\begin{vmatrix} b_{M1}^{(1)} & b_{M2}^{(1)} \\ b_{s1}^{(1)} & b_{s2}^{(1)} \end{vmatrix}}{b_{s2}^{(1)}} & -\frac{b_{M2}^{(1)}}{b_{s2}^{(1)}}
\end{pmatrix}$$

In $\mathbf{A}^{(5)}$, use $b_{ij}^{(2)}$ to denote the element in $\mathbf{B}^{(5)}$, then we have:

$$b_{ij}^{(2)} < 0, \text{for i} = \text{s}, \ \text{s} + 1, \ \cdots, \ \text{M}, \ \text{and j} = 1, 2.$$

Similarly, we can move all the negative row vectors in $\mathbf{B}^{(5)}$ to the bottom side of it while guaranteeing the identity matrix structure in the front. We note that the number of those negative vectors are greater than P. Repeat Step 1∼5 till for $i =$

1, 2, $\cdots$, M, we have:

$$b_{i1}b_{i2} \geq 0, \tag{4.2.42}$$

i.e., there is no more row vector whose elements possess different sign. Thus the row vector of $\mathbf{B}$ is either non-positive or non-negative.

For the second case, i.e., if $\mathbf{A}$ has zero-cover, then $\mathbf{B}$ contains one negative column, or one non-positive column for which the position of the zero is occupied by a negative quantity in another column. Similar to the first case, if the elements in some row of $\mathbf{B}$ have different signs, then we can always transform this kind of row to the negative one. Finally, $\mathbf{B}$ only contains two non-positive column vector and there is no zero row vector in $\mathbf{B}$.

This concludes the proof of Lemma 4.2.2 □

4.2.3 Échelon Form of Low-Rank Matrices

Low-rank matrices generally possess some interesting échelon forms. In the following, we will examine the structure of the matrix with its rank equals to 2 in more detail.

Theorem 4.2.2. *Suppose* $\mathbf{A}$ *is an* $M \times N$ *real matrix such that* $R_r(\mathbf{A}) = 2$*. If* $\mathbf{A}$ *is full cover, then* $\mathbf{A}$ *can be transformed into:*

$$\mathbf{A} \rightarrow \begin{pmatrix} \mathbf{I}_2 & \mathbf{B}_{2\times(N-2)} \\ \mathbf{0}_{(M-2)\times 2} & \mathbf{0}_{(M-2)\times(N-2)} \end{pmatrix} \tag{4.2.43}$$

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

Proof. Without loss of generality, given an $M \times N$ real matrix $\mathbf{A}$ and suppose it has full rank and its rank is 2, i.e., $M = 2$. If $\mathbf{A}$ has full-cover, then by Theorem 4.2.1, $\mathbf{A}$ can be transformed into:

$$\begin{pmatrix} 1 & 0 & b_{11} & b_{12} & \cdots & b_{1,N-2} \\ 0 & 1 & b_{21} & b_{22} & \cdots & b_{2,N-2} \end{pmatrix} \tag{4.2.44}$$

where $b_{11}, b_{12}, \cdots, b_{1,N-2} > 0$. If there exist $b_{2i} < 0$, $i \in \{1, \cdots, N-2\}$, then let

$$t = \max_{i \in \{1,\ \cdots,\ N-2\}} \{-\frac{b_{2i}}{b_{1i}}, b_{2i} < 0\} = -\frac{b_{2j}}{b_{1j}} \tag{4.2.45}$$

Multiplying the first row of $\mathbf{A}$ by t and adding the product to the second row of it. Then we have:

$$\begin{pmatrix} 1 & 0 & b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1,N-2} \\ t & 1 & b_{21} - b_{11}\frac{b_{2j}}{b_{1j}} & b_{22} - b_{12}\frac{b_{2j}}{b_{1j}} & \cdots & 0 & \cdots & b_{2,N-2} - b_{1,N-2}\frac{b_{2j}}{b_{1j}} \end{pmatrix} \tag{4.2.46}$$

In the next step, multiplying the first row of above matrix by $\frac{1}{b_{1j}}$ and exchanging the first column with the j-th column so that the identity matrix structure in the left-hand side part can be guaranteed, we have:

$$\begin{pmatrix} 1 & 0 & \frac{b_{11}}{b_{1j}} & \frac{b_{12}}{b_{1j}} & \cdots & \frac{1}{b_{1j}} & \cdots & \frac{b_{1,N-2}}{b_{1j}} \\ 0 & 1 & b_{21} - b_{11}\frac{b_{2j}}{b_{1j}} & b_{22} - b_{12}\frac{b_{2j}}{b_{1j}} & \cdots & -\frac{b_{2j}}{b_{1j}} & \cdots & b_{2,N-2} - b_{1,N-2}\frac{b_{2j}}{b_{1j}} \end{pmatrix} \tag{4.2.47}$$

Now, for $b_{2i} < 0, \quad i \in \{1,\ \cdots,\ N-2\}$,

$$b_{2i} - b_{1i}\frac{b_{2j}}{b_{1j}} = b_{1i}(\frac{b_{2i}}{b_{1i}} - \frac{b_{2j}}{b_{1j}}) > 0 \tag{4.2.48}$$

And, for $b_{2i} > 0$, it is obvious that $b_{2i} - b_{1i}\frac{b_{2j}}{b_{1j}} > 0$. Thus, after these two steps, we obtain a $\mathbf{B}$ which has two non-negative row vectors. □

Theorem 4.2.3. *Suppose* $\mathbf{A}$ *is an* $M \times N$ *real matrix and* $R_r(\mathbf{A}) = 2$*. Then* $\mathbf{A}$ *is zero cover if and only if it can be transformed into the form:*

$$\mathbf{EAP} = \begin{pmatrix} \mathbf{I}_2 & \mathbf{B}^+ & \mathbf{B}^- \\ \mathbf{0}_{(N-2)\times 2} & \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{4.2.49}$$

where all the elements in $\mathbf{B}^+$ *are non-negative and the elements in* $\mathbf{B}^-$ *are all non-positive. 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 $\mathbf{A}$ has full rank, i.e., $M = 2$. Firstly, let us proof the sufficient condition. Given a zero-cover matrix $\mathbf{A}$, then by Theorem 4.2.1, it can be transformed into:

$$\begin{pmatrix} 1 & 0 & b_{11} & b_{12} & \cdots & b_{1,N-2} \\ 0 & 1 & b_{21} & b_{22} & \cdots & b_{2,N-2} \end{pmatrix} \tag{4.2.50}$$

then let:

$$t = \max_{i\in\{1,\,\cdots,\,N-2\}} \{-\frac{b_{2i}}{b_{1i}},\ b_{1i}b_{2i} < 0\} = -\frac{b_{2j}}{b_{1j}} \tag{4.2.51}$$

Multiplying the first row by t and adding the product to the second row. Then we have:

$$\begin{pmatrix} 1 & 0 & b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1,N-2} \\ t & 1 & b_{21}-b_{11}\frac{b_{2j}}{b_{1j}} & b_{22}-b_{12}\frac{b_{2j}}{b_{1j}} & \cdots & 0 & \cdots & b_{2,N-2}-b_{1,N-2}\frac{b_{2j}}{b_{1j}} \end{pmatrix} \tag{4.2.52}$$

In the next step, multiplying the first row of above matrix with $\frac{1}{b_{1j}}$ and exchanging the first column with the j-th column so that the identity matrix structure in the left-hand side part can be guaranteed. Then we will have:

$$\begin{pmatrix} 1 & 0 & \frac{b_{11}}{b_{1j}} & \frac{b_{12}}{b_{1j}} & \cdots & \frac{1}{b_{1j}} & \cdots & \frac{b_{1,N-2}}{b_{1j}} \\ 0 & 1 & b_{21}-b_{11}\frac{b_{2j}}{b_{1j}} & b_{22}-b_{12}\frac{b_{2j}}{b_{1j}} & \cdots & -\frac{b_{2j}}{b_{1j}} & \cdots & b_{2,N-2}-b_{1,N-2}\frac{b_{2j}}{b_{1j}} \end{pmatrix} \tag{4.2.53}$$

For those $b_{1i} > 0$, $i \in \{1, \ \cdots, \ N-2\}$, we have:

$$b_{2i} - b_{1i}\frac{b_{2j}}{b_{1j}} = b_{1i}(\frac{b_{2i}}{b_{1i}} - \frac{b_{2j}}{b_{1j}}) > 0 \tag{4.2.54}$$

For those $b_{1i} < 0$, $i \in \{1, \ \cdots, \ N-2\}$, if $b_{2i} < 0$, then we have:

$$b_{2i} - b_{1i}\frac{b_{2j}}{b_{1j}} < 0 \tag{4.2.55}$$

If $b_{2i} > 0$, the sign of $(b_{2i} - b_{1i}\frac{b_{2j}}{b_{1j}})$ is uncertain. After above steps, and by doing some certain column permutations, the matrix $\mathbf{A}$ can be transformed into the following form:

$$\begin{pmatrix} \mathbf{I} & \mathbf{B}_{(1)} & \mathbf{B}_{(2)} & \mathbf{B}_{(3)} \end{pmatrix} \tag{4.2.56}$$

where $\mathbf{B}_{(1)}$ is a non-negative matrix, $\mathbf{B}_{(3)}$ is a non-positive 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 above matrix as:

$$\begin{pmatrix} 1 & 0 & b_{11}^{(1)} & b_{12}^{(1)} & \cdots & b_{1,N-2}^{(1)} \\ 0 & 1 & b_{21}^{(1)} & b_{22}^{(1)} & \cdots & b_{2,N-2}^{(1)} \end{pmatrix} \tag{4.2.57}$$

Then let:

$$m = \max_{i\in\{1,\ \cdots,\ N-2\}} \{-\frac{b_{1i}^{(1)}}{b_{2i}^{(1)}},\ b_{2i}^{(1)} b_{1i}^{(1)} < 0\} = -\frac{b_{1s}^{(1)}}{b_{2s}^{(1)}} \tag{4.2.58}$$

Multiplying the second row of above matrix by m and adding the product to the first row. Then we have:

$$\begin{pmatrix} 1 & m & b_{11}^{(1)} - b_{21}^{(1)}\frac{b_{1s}^{(1)}}{b_{2s}^{(1)}} & b_{12}^{(1)} - b_{22}^{(1)}\frac{b_{1s}^{(1)}}{b_{2s}^{(1)}} & \cdots & 0 & \cdots & b_{1,N-2}^{(1)} - b_{2,N-2}^{(1)}\frac{b_{1s}^{(1)}}{b_{2s}^{(1)}} \\ 0 & 1 & b_{21}^{(1)} & b_{22}^{(1)} & \cdots & b_{2s}^{(1)} & \cdots & b_{2,N-2}^{(1)} \end{pmatrix} \tag{4.2.59}$$

Multiplying the second row of above matrix with $\frac{1}{b_{2s}^{(1)}}$ and exchanging the second column with the s-th column so that the identity matrix structure in the left-hand side part can be guaranteed, we have:

$$\begin{pmatrix} 1 & 0 & b_{11}^{(1)} - b_{21}^{(1)}\frac{b_{1s}^{(1)}}{b_{2s}^{(1)}} & b_{12}^{(1)} - b_{22}^{(1)}\frac{b_{1s}^{(1)}}{b_{2s}^{(1)}} & \cdots & -\frac{b_{1s}^{(1)}}{b_{2s}^{(1)}} & \cdots & b_{1,N-2}^{(1)} - b_{2,N-2}^{(1)}\frac{b_{1s}^{(1)}}{b_{2s}^{(1)}} \\ 0 & 1 & \frac{b_{21}^{(1)}}{b_{2s}^{(1)}} & \frac{b_{22}^{(1)}}{b_{2s}^{(1)}} & \cdots & \frac{1}{b_{2s}^{(1)}} & \cdots & \frac{b_{2,N-1}^{(1)}}{b_{2s}^{(1)}} \end{pmatrix} \tag{4.2.60}$$

In the above matrix, we have if $b_{2i}^{(1)} > 0$, $i \in \{1,\ \cdots,\ N-2\}$, then

$$b_{1i}^{(1)} - b_{2i}^{(1)}\frac{b_{1s}^{(1)}}{b_{2s}^{(1)}} > 0 \tag{4.2.61}$$

and if $b_{2i}^{(1)} < 0$, $i \in \{1, \cdots, N-2\}$, then

$$b_{1i}^{(1)} - b_{2i}^{(1)} \frac{b_{1s}^{(1)}}{b_{2s}^{(1)}} < 0 \tag{4.2.62}$$

As a result, if $\mathbf{A}$ has zero cover, then it can be transformed into the right hand side form in Eq. (4.2.49).

To prove the necessity: Suppose $\mathbf{A}$ can be transformed into the following form:

$$\begin{pmatrix} \mathbf{I}_2 & \mathbf{B}^+ & \mathbf{B}^- \end{pmatrix} \tag{4.2.63}$$

Consider the case when $\mathbf{B}^-$ contains two non-negative vectors with their negative terms lie in different rows. Then $\mathbf{Ax} = \mathbf{0}$ can be written as:

$$\begin{pmatrix} \mathbf{I}_2 & \mathbf{B}^+ & \bar{\mathbf{B}}^- & \mathbf{b_1} & \mathbf{b_2} \end{pmatrix} \times \begin{pmatrix} x_1 \\ x_2 \\ \mathbf{x}^+ \\ \mathbf{x}^- \\ x \\ y \end{pmatrix} = \mathbf{0} \tag{4.2.64}$$

where

$$\mathbf{b_1} = \begin{pmatrix} b_1 \\ 0 \end{pmatrix}, \mathbf{b_2} = \begin{pmatrix} 0 \\ b_2 \end{pmatrix} \tag{4.2.65}$$

and b_1, b_2 are both negative, $\bar{\mathbf{B}}^-$ is the matrix formed by deleting the column vectors

$\mathbf{b_1}$ and $\mathbf{b_2}$ from $\mathbf{B}^-$. Then we have:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} - \mathbf{B}^+\mathbf{x}^+ - \bar{\mathbf{B}}^-\mathbf{x}^- \tag{4.2.66}$$

We can let the elements in $\mathbf{x}^+$ and $\mathbf{x}^-$ be any positive value. If we let x and y be positive and large enough, we can still obtain positive x_1 and x_2. In this case, all elements in $\mathbf{x}$ are positive and satisfy the equation $\mathbf{Ax} = \mathbf{0}$. By Theorem 3.1.1, $\mathbf{A}$ has zero cover.

This concludes the proof of Theorem 4.2.3. □

4.2.4 The Analysis of the Linear Equations System with Nonnegativity Constraints on Solutions through the Échelon Structure

From the échelon structure of the matrix, we can gain a deeper understanding of the system of linear equations having nonnegativity constraints on variables: Given an $M \times N$ real matrix $\mathbf{A}$ and an $M \times 1$ column vector $\mathbf{b}$. Consider the linear equations system with nonnegativity constraints on variable $\mathbf{x}$: $\mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$. According to the above discussion, we have the cover order of $\mathbf{A}$ given by

$$R_c(\mathbf{A}) = \sum_{i=1}^{s} N_i + s \tag{4.2.67}$$

where s is obtained from échelon transformation, and the number of uncovered variable x_i in $\mathbf{x}$ or the number of uncovered columns in the matrix $\mathbf{A}$ is

$$\bar{R}_c(\mathbf{A}) = N - R_c(\mathbf{A}) \tag{4.2.68}$$

Then, we obtain the following result:

Lemma 4.2.3. *Let* $\mathbf{A} \in \mathbb{R}^{M\times N}$ *and* $\mathbf{b} \in \mathbb{R}^{M\times 1}$. *For non-homogeneous linear equations with the nonnegativity constraints on variable* $\mathbf{x}$*:* $\mathbf{A}\mathbf{x} = \mathbf{b}\ ,\ \mathbf{x} \geq \mathbf{0}$.

1. *The system has a unique solution in* $\mathbb{R}_N^+$ *if and only if* $R_r(\mathbf{A}) - s = N - R_c(\mathbf{A})$ *and* $R_c(\mathbf{A}) = R_c(\tilde{\mathbf{A}})$*, where* $\tilde{\mathbf{A}} = (\mathbf{A}, -\mathbf{b})$ *and* s *is the number of transformation in the structure arrangement step of échelon transformation..*

2. *The system has infinite solution in* $\mathbb{R}_N^+$ *when* $R_r(\mathbf{A}) - s < N - R_c(\mathbf{A})$ *and* $R_c(\tilde{\mathbf{A}}) \leq R_c(\mathbf{A})$.

■

By reviewing Theorem 3.2.1 and Theorem 3.2.2, it is apparent that the result obtained in Lemma 4.2.3 is consistent with the conclusion that we obtained in Chapter 2.

4.3 Non-Negatively Linear Independence

Given an $M \times N$ real matrix $\mathbf{A}$, suppose $\mathbf{A}$ has full rank and $M < N$. Consider the case when $\mathbf{A}$ does not have full cover, i.e., $R_c(\mathbf{A}) < N$. Once we have its échelon form, we can divide this form into several components:

$$\mathbf{EAP} = \begin{pmatrix} \mathbf{I}_s & \mathbf{0} & \mathbf{B^{(1)}} & \mathbf{B^{(3)}} \\ \mathbf{0} & \mathbf{I}_{M-s} & \mathbf{B^{(2)}} & \mathbf{B^{(4)}} \end{pmatrix} \tag{4.3.1}$$

where in Eq. (4.3.1), $\mathbf{B^{(1)}}$ is a $s \times (\sum_{i=1}^{s} N_i)$ real matrix, the elements in $\mathbf{B^{(3)}}$ are all zeros, $\mathbf{B^{(4)}}$ is a $(M-s) \times (N - M - \sum_{i=1}^{s} N_i)$ real matrix and is $\bar{\mathbf{B}}$ in Eq. (4.2.5) exactly. Generally, we can denote the matrix formed by the block matrices $\mathbf{I}_s$ and $\mathbf{B^{(1)}}$ as:

$$\mathbf{A}^u = \begin{pmatrix} \mathbf{I}_s & \mathbf{B^{(1)}} \end{pmatrix} \tag{4.3.2}$$

It is observed that $\mathbf{A}^u$ has full cover. The matrix formed by the lower side block matrices: $\mathbf{I}_{(M-s)}$, $\mathbf{B^{(2)}}$ and $\mathbf{B^{(4)}}$ in Eq. (4.3.1) is denoted as:

$$\mathbf{A}^l = \begin{pmatrix} \mathbf{0}_{(M-s)\times s} & \mathbf{I}_{(M-s)} & \mathbf{B^{(2)}} & \mathbf{B^{(4)}} \end{pmatrix} \tag{4.3.3}$$

Note that the cover order of $\mathbf{A}^l$ is zero. As a result, the cover order of the matrix $\mathbf{A}$ will not be affected by deleting the row vectors from this part. There is, however, a limit of $M - s$ in the amount of rows that can be deleted, which means that, if we continue to delete the row vectors after deleting these rows, the cover order of $\mathbf{A}$ will be altered. In the following, we first give the definition of non-negatively linear dependence and non-negatively linear independence. Then, according to the échelon form and Theorem 3.1.1, we define the number of the non-negatively linear independent row vectors and non-negatively linear dependent column vectors in the matrix $\mathbf{A}$.

Definition 4.3.1. *A sequence of vectors* $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ *from a vector space* $\mathbb{V}$ *is said to be non-negatively linear dependent, if there exist non-negative scalars* $x_1,\ x_2,\ \cdots,\ x_n$*, not all zero, such that*

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{0} \tag{4.3.4}$$

where $\mathbf{0}$ *denotes the zero vector.* ■

Definition 4.3.2. *A sequence of vectors* $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ *is said to be non-negatively linear independent if it is not non-negatively linear dependent, that is, if for non-negative scalars* $x_1,\ x_2,\ \cdots,\ x_n$*, the equation*

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{0} \tag{4.3.5}$$

can only be satisfied by $x_i = 0$ *for* $i = 1,\ 2,\ \cdots,\ n$*.* ■

With the definition of non-negatively linear independent, given a real matrix, we define the order of non-negatively linear independence as follows.

Definition 4.3.3. *Let* $\mathbf{A}$ *be an* $M \times N$ *real matrix, then* $\mathbf{A}$ *is called non-negatively linear independent with order* s *if the columns in every* s *rows are non-negatively linear independent, but the columns in* $s-1$ *rows are not fully non-negatively linear independent, where* $s \leq M$*.* ■

Definition 4.3.4. *Let* $\mathbf{A}$ *be an* $M \times N$ *real matrix, then* $\mathbf{A}$ *is called non-negatively linear dependent with order* s *if every* s *columns are non-negatively linear dependent, where* $s \leq N$*.* ■

Notice that in the above two definitions, we are more concerned with the number of rows and columns than the exact position. Furthermore, the definitions above provide us with a new perspective on what the cover order means as stated in the following lemma.

Lemma 4.3.1. *The cover order of the matrix* $\mathbf{A}$ *is the number of the non-zero columns of the sub-matrix consisting of the minimum number of row vectors that the columns in it are non-negatively linear independent.* ■

4.4 Dual Relationship between Cover and Uncover

In this section, we will first introduce the construction of the generalized inverse of any given matrix $\mathbf{A}$ (4; 11; 44; 49), from which we can express the solutions to the system of linear equations $\mathbf{Ax} = \mathbf{b}$ (30). Then with the aid of Farkas' Lemma, we derive an interesting dual property between a column vector being covered in $\mathbf{A}$ and an associated column vector being uncovered in the corresponding matrix $\tilde{\mathbf{A}}$.

Definition 4.4.1. *Let* $\mathbf{A}$ *be an* $M \times N$ *real matrix. Then, an* $N \times M$ *real matrix* $\mathbf{G}$ *is said to be a generalized inverse of* $\mathbf{A}$ *if:*

$$\mathbf{AGA} = \mathbf{A} \tag{4.4.1}$$

. ■

Lemma 4.4.1. *Let* $N \times M$ *real matrix* $\mathbf{G}_0$ *be a generalized inverse of an* $M \times N$ *real matrix* $\mathbf{A}$. *Then, all the generalized inverse matrices of the matrix* $\mathbf{A}$ *can be* *represented by*

$$\mathbf{G} = \mathbf{G}_0 - \mathbf{X} + \mathbf{G}_0\mathbf{A}\boldsymbol{\Omega}\mathbf{A}\mathbf{G}_0, \tag{4.4.2}$$

where $\boldsymbol{\Omega}$ *is an arbitrarily given* $N \times M$ *real matrix.* ■

Lemma 4.4.2. *Let* $\mathbf{A}$ *be an* $M \times N$ *real matrix and* $\mathbf{A} \neq \mathbf{0}$. *Then, a specific generalized inverse of* $\mathbf{A}$ *is given by*

$$\mathbf{G}_0 = (\mathbf{A}^T\mathbf{A})^\dagger \mathbf{A}^T, \tag{4.4.3}$$

where

$$(\mathbf{A}^T\mathbf{A})^\dagger = \mathbf{V} \begin{pmatrix} \boldsymbol{\Lambda}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^T \tag{4.4.4}$$

with the singular value decomposition (SVD) of $\mathbf{A}$ *being given by*

$$\mathbf{A} = \mathbf{U} \begin{pmatrix} \boldsymbol{\Lambda}^{\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^T \tag{4.4.5}$$

■

With this specific generalized inverse matrix, all the generalized inverse matrices of $\mathbf{A}$ can be expressed by

$$\mathbf{G} = (\mathbf{A}^T\mathbf{A})^\dagger \mathbf{A}^T - \boldsymbol{\Omega} + (\mathbf{A}^T\mathbf{A})^\dagger \mathbf{A}^T\mathbf{A}\boldsymbol{\Omega}\mathbf{A}(\mathbf{A}^T\mathbf{A})^\dagger \mathbf{A}^T, \tag{4.4.6}$$

where $\boldsymbol{\Omega}$ is an arbitrarily given $N \times M$ real matrix.

Using the generalized inverse matrices, we can express all the solutions to the system of linear equations: $\mathbf{Ax} = \mathbf{b}$, as

$$\mathbf{x} = \mathbf{Gb}. \tag{4.4.7}$$

Thus, we have the following lemma to uniformly express all solutions to the linear equation system.

Lemma 4.4.3. *Let $\mathbf{G}_0$ be a generalized inverse of an $M \times N$ matrix $\mathbf{A}$. Then, all the solutions to a system of linear equations: $\mathbf{Ax} = \mathbf{b}$ are given by*

$$\mathbf{x} = \mathbf{G}_0\mathbf{b} + (\mathbf{G}_0\mathbf{A} - \mathbf{I})\mathbf{z}, \tag{4.4.8}$$

where $\mathbf{z}$ is an arbitrary vector in $\mathbb{R}^N$. ■

Thus, in the derivation of the following Theorem 4.4.1, with the aid of Lemma 4.4.3, we are able to write a general expression of the solutions for the corresponding linear equations.

Lemma 4.4.4 (Farkas' Lemma (variant))**.** *Let $\mathbf{M}$ be an $M \times N$ real matrix and $\mathbf{b} \in \mathbb{R}^M$. Then either:*

1. *There is an $\mathbf{x} \in \mathbb{R}^N$ such that $\mathbf{Mx} \leq \mathbf{b}$; or*

2. *There is a $\mathbf{y} \in \mathbb{R}^M$ such that $\mathbf{y} \geq \mathbf{0}, \mathbf{M}^T\mathbf{y} = \mathbf{0}$ and $\mathbf{b}^T\mathbf{y} < 0$.*

■

With Farkas' Lemma, we are able to obtain the following dual property of a vector being covered and uncovered:

Theorem 4.4.1. *Let* $\mathbf{A} = \begin{pmatrix} \mathbf{I} & \mathbf{B} \end{pmatrix}$, *where* $\mathbf{A} \in \mathbb{R}^{M\times N}$, $\mathbf{I} = (\mathbf{e}_1, \mathbf{e}_2, \ \cdots, \ \mathbf{e}_M) \in \mathbb{R}^{M\times M}$ *and* $\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \ \cdots, \ \mathbf{b}_{N-M}) \in \mathbb{R}^{M\times(N-M)}$. *Then,* $\mathbf{b}_k$ *is uncovered in* $\mathbf{A}$ *if and only if* $\mathbf{e}_k$ *is covered in the matrix* $\tilde{\mathbf{A}}$, *for* $k = 1, 2, \ \cdots, \ N - M$, *where* $\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{I}_{N-M} & -\mathbf{B}^T \end{pmatrix}$.

■

Proof. When we consider a specific matrix $\mathbf{A} = \begin{pmatrix} \mathbf{I} & \mathbf{B} \end{pmatrix}$, where $\mathbf{A} \in \mathbb{R}^{M\times N}$, $\mathbf{I} \in \mathbb{R}^{M\times M}$ and $\mathbf{B} \in \mathbb{R}^{M\times(N-M)}$. Let $K = N - M$ and for discussion convenience, we write $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_N)$, and the sub-matrix $\bar{\mathbf{A}}_N = (\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_{N-1})$ by deleting the N-the column from the matrix $\mathbf{A}$. We also write $\mathbf{B}$ as $(\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_K)$ and the sub-matrix $\bar{\mathbf{B}}_K = (\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_{K-1})$ is obtained by deleting the K-th column of the matrix $\mathbf{B}$.

Then, according to the equation in Definition 4.4.1, we see right away that a generalized inverse matrix of $\bar{\mathbf{A}}_N$ is given by:

$$\mathbf{G}_0 = \begin{pmatrix} \mathbf{I} & \mathbf{0} \end{pmatrix}^T \tag{4.4.9}$$

where $\mathbf{G}_0 \in \mathbb{R}^{(N-1)\times M}$ and the identity matrix $\mathbf{I} \in \mathbb{R}^{M\times M}$ in above equation. Hence, according to Lemma 4.4.3, all solutions to the system of linear equations: $\bar{\mathbf{A}}_N\mathbf{x} = -\mathbf{b}_K$ can be generated by

$$\mathbf{x} = -\mathbf{G}_0\mathbf{b}_K + (\mathbf{I} - \mathbf{G}_0\bar{\mathbf{A}}_N)\mathbf{z} \tag{4.4.10}$$

Therefore, according to Theorem 3.2.1, we have: the system of linear equations: $\bar{\mathbf{A}}_N\mathbf{x} = -\mathbf{b}_K$ has a solution in $\mathbb{R}_+^{N-1}$ if and only if the cover order of the matrix

$\begin{pmatrix}\bar{\mathbf{A}}_N & \mathbf{b}_K\end{pmatrix}$, i.e., the cover order of the matrix $\mathbf{A}$ is less than or equal to the cover order of $\bar{\mathbf{A}}_N$. Thus by adding $\mathbf{a}_N$ or $\mathbf{b}_K$ to the right hand side of $\bar{\mathbf{A}}_N$, and if $\mathbf{a}_N$ or $\mathbf{b}_K$ is not covered in $\begin{pmatrix}\bar{\mathbf{A}}_N & \mathbf{b}_K\end{pmatrix}$, then the cover order of $\begin{pmatrix}\bar{\mathbf{A}}_N & \mathbf{b}_K\end{pmatrix}$ will not be greater than the cover order of $\bar{\mathbf{A}}_N$, and hence, the system of linear equation $\bar{\mathbf{A}}_N\mathbf{x} = -\mathbf{b}_K$ will have a solution in $\mathbb{R}_+^{N-1}$.

From the above, we establish that: the column vector $\mathbf{a}_N$ or $\mathbf{b}_K$ is not covered in $\mathbf{A}$ is equivalent to that there exists a vector $\mathbf{z}_0 \in \mathbb{R}^{N-1}$ satisfying:

$$-\mathbf{G}_0\mathbf{b}_K + (\mathbf{I} - \mathbf{G}_0\bar{\mathbf{A}}_N)\mathbf{z}_0 \geq \mathbf{0}, \tag{4.4.11}$$

i.e., $(\mathbf{G}_0\bar{\mathbf{A}}_N - \mathbf{I})\mathbf{z}_0 \leq -\mathbf{G}_0\mathbf{b}_K$. However, from Farkas' lemma, letting

$$\begin{aligned} \mathbf{M} &= \mathbf{G}_0\bar{\mathbf{A}}_N - \mathbf{I}, & (4.4.12a)\\ \mathbf{b} &= -\mathbf{G}_0\mathbf{b}_K, & (4.4.12b) \end{aligned}$$

we have *either* the inequality $(\mathbf{G}_0\bar{\mathbf{A}}_N - \mathbf{I})\mathbf{z}_0 \leq -\mathbf{G}_0\mathbf{b}_K$ holds, *or* there exists $\mathbf{y} \in \mathbb{R}_+^{N-1}$ such that

$$\begin{aligned} (\mathbf{G}_0\bar{\mathbf{A}}_N - \mathbf{I})^T\mathbf{y} &= \mathbf{0}, & (4.4.13a)\\ -\mathbf{b}_K^T\mathbf{G}_0^T\mathbf{y} &< \mathbf{0}. & (4.4.13b) \end{aligned}$$

Thus, we infer that: $\mathbf{a}_N$ or $\mathbf{b}_K$ is covered in $\mathbf{A}$ if and only if there exists $\mathbf{y} \in \mathbb{R}_+^{N-1}$ which satisfies Eqs. (4.4.13). Now, let us examine what it means when such a vector $\mathbf{y}$ exists.

Rewrite $\mathbf{y}$ as $\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}$, where $\mathbf{y}_1 \in \mathbb{R}_+^M$ and $\mathbf{y}_2 \in \mathbb{R}_+^{N-M-1}$, then Eqs. (4.4.13) can be written as:

$$\mathbf{y}_2 = \bar{\mathbf{B}}_K^T \mathbf{y}_1, \tag{4.4.14a}$$

$$\mathbf{b}_K^T \mathbf{y}_1 > 0. \tag{4.4.14b}$$

Now, consider the following $(N - M) \times N$ real matrix:

$$\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{I} & \mathbf{0} & -\bar{\mathbf{B}}_K^T \\ \mathbf{0} & 1 & -\mathbf{b}_K^T \end{pmatrix}. \tag{4.4.15}$$

The size of the identity matrix $\mathbf{I}$ in the above matrix is $(N - M - 1) \times (N - M - 1)$ and now let:

$$\tilde{\mathbf{x}} = \begin{pmatrix} \mathbf{y}_2 \\ y_3 \\ \mathbf{y}_1 \end{pmatrix} \tag{4.4.16}$$

Then according to Eq.(4.4.14a) and inequality (4.4.14b), we have:

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} & -\bar{\mathbf{B}}_K^T \\ \mathbf{0} & 1 & -\mathbf{b}_K^T \end{pmatrix} \begin{pmatrix} \mathbf{y}_2 \\ y_3 \\ \mathbf{y}_1 \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ y_3 - \mathbf{b}_K^T \mathbf{y}_1 \end{pmatrix} \tag{4.4.17}$$

For the homogeneous linear equations system: $\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \mathbf{0}$, we have

$$y_3 = \mathbf{b}_K^T \mathbf{y}_1 > 0 \tag{4.4.18}$$

Thus we say that there exists a non-negative vector $\tilde{\mathbf{x}} \in \mathbb{R}_+^N$ with $\tilde{\mathbf{x}} \neq \mathbf{0}$ such that $\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \mathbf{0}$ holds, where y_3 in $\tilde{\mathbf{x}}$ is positive. By Theorem 3.1.1, we see that the column vector corresponding to the variable y_3 in the product $\tilde{\mathbf{A}}\tilde{\mathbf{x}}$ is not covered. We can rewrite the matrix $\tilde{\mathbf{A}}$ as:

$$\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{I} & -\mathbf{B}^T \end{pmatrix} = \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_K & -\mathbf{B}^T \end{pmatrix} \tag{4.4.19}$$

We observe from Eq. (4.4.17) that the column vector $\mathbf{e}_K \in \mathbb{R}^{N-M}$ corresponds to the variable y_3 in the product $\tilde{\mathbf{A}}\tilde{\mathbf{x}}$. Thus we have, $\mathbf{e}_K$ is not covered in the matrix $\begin{pmatrix} \mathbf{I}_{N-M} & -\mathbf{B}^T \end{pmatrix}$. This establishes the statement: "$\mathbf{e}_K$ is not covered in the matrix $\begin{pmatrix} \mathbf{I}_{N-M} & -\mathbf{B}^T \end{pmatrix}$ if and only if there exists $\mathbf{y} \in \mathbb{R}_+^{N-1}$ satisfying Eqs. (4.4.13) " which, in turn, implies and is implied by: "$\mathbf{a}_N$ or $\mathbf{b}_K$ is covered in the matrix $\mathbf{A}$".

Conversely, $\mathbf{e}_K$ is covered in the matrix $\begin{pmatrix} \mathbf{I}_{N-M} & -\mathbf{B}^T \end{pmatrix}$ iff $\mathbf{a}_N$ or $\mathbf{b}_K$ is not covered in $\mathbf{A}$. This completes the proof. □

Remarks: Theorem 4.4.1 establishes a dual relationship for the coverage of vectors in the matrix $\mathbf{A}$ and in the matrix $\tilde{\mathbf{A}}$. If the coverage condition is not obvious in $\mathbf{A}$, it may be clearer when the dual matrix $\tilde{\mathbf{A}}$ is inspected.

4.5 Inner Hyper-Rectangle Cover

Our main task in this section is to reveal the properties of inner hyper-rectangle in the set

$$\left\{\mathbf{x} : 0 \le \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \le \tau^2, \mathbf{x} \in \mathbb{R}_+^N\right\}. \tag{4.5.1}$$

It is shown that these properties are closely related to zero-cover. To make our idea more understandable, we revisit Example 2.1.3. By injecting new insight into Fig. 2.3, we come up with Fig. 4.1 from which it can be observed that inside the domain determined by

$$\{(x_1, x_2) : (x_1 - x_2)^2 \le \tau^2, x_1, x_2 \ge 0\}, \tag{4.5.2}$$

there is a largest inner square with four vertices being $(0,0), (0,1), (1,0)$ and $(1,1)$. More importantly, we observe that shifting this square in the direction $\xi(1,1)$ for any positive number ξ gives us a new square which is still inside $\{(x_1, x_2) : (x_1 - x_2)^2 \le \tau^2, x_1, x_2 \ge 0\}$. We now formally state this important observation in a general case:

Theorem 4.5.1. *If the cover order of an $M \times N$ real matrix $\mathbf{A}$ is R_c and $0 \le R_c < N$, then, the following two statements are true.*

1. *There exists a nonnegative vector $\mathbf{v} \in \mathbb{R}_+^N$ such that*

$$\begin{aligned}&\left\{\mathbf{x} + \xi\mathbf{v} : 0 \le x_i \le \frac{\tau}{\sqrt{\lambda_{\max}}}, 1 \le i \le N, \xi > 0\right\}\\ &\subseteq \left\{\mathbf{x} : 0 \le \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \le \tau^2, \mathbf{x} \in \mathbb{R}_+^N\right\}\end{aligned} \tag{4.5.3}$$

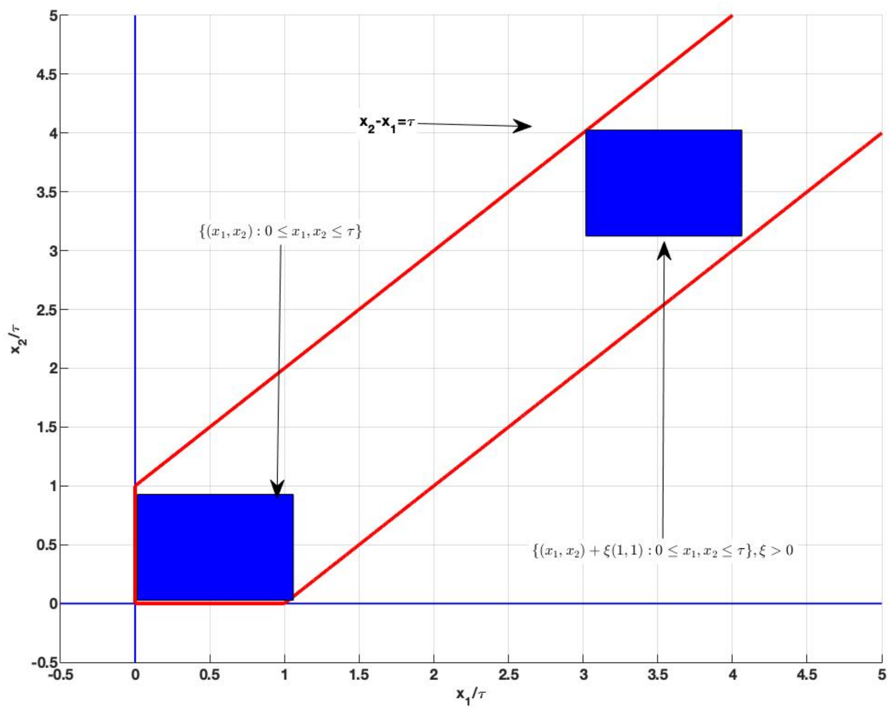

Figure 4.1: Example of a zero-cover 2×2 PSD matrix.

holds for any given positive constant τ, where $\lambda_{\max}$ denotes the maximum eigenvalue of $\mathbf{A}^T\mathbf{A}$.

2. *The subprincipal matrix of $\mathbf{A}^T\mathbf{A}$, which is formed by the rows and columns having the same indices as the positive entries of $\mathbf{v}$, is zero-cover.*

■

Proof. Proof of Statement 1): Let $\lambda_{\max}$ denotes the maximum eigenvalue of $\mathbf{A}^T\mathbf{A}$, then we have:

$$\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \lambda_{\max}\|\mathbf{x}\|_2^2, \tag{4.5.4}$$

If $\lambda_{\max}\|\mathbf{x}\|_2^2 \leq \tau^2$ for any positive τ, then $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2$ holds. This result gives us that

$$\begin{aligned}
&\left\{\mathbf{x} : 0 \leq x_i \leq \frac{\tau}{\sqrt{\lambda_{\max}}}, 1 \leq i \leq N, \xi > 0\right\} \\
&\subseteq \left\{\mathbf{x} : \lambda_{\max}\|\mathbf{x}\|_2^2 \leq \tau^2, \mathbf{x} \in \mathbb{R}_+^N\right\} \\
&\subseteq \left\{\mathbf{x} : 0 \leq \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2, \mathbf{x} \in \mathbb{R}_+^N\right\}
\end{aligned} \tag{4.5.5}$$

In the following, we construct the desired nonnegative vector $\mathbf{v}$. By Theorem 4.1.1, we know that if the cover order of $\mathbf{A}$ is R_c, then, there exists a nonnegative vector $\mathbf{p}$ with R_c positive entries and $(N - R_c)$ zero-valued entries. Denote the indices of these positive entries of $\mathbf{p}$ by $\ell_i^{(+)}$ and the indexes of the zero-valued entries by $\ell_k^{(0)}$, where $i = 1, \cdots, R_c$ and $k = 1, \cdots, (N - R_c)$. Then, similar to the proof of Theorem 4.1.1, we form an $(N - R_c) \times (N - R_c)$ sub-matrix of $\mathbf{A}^T\mathbf{A}$ by using the columns of $\mathbf{A}$ indexed by $\ell_i^{(0)}, i = 1, \cdots, (N - R_c)$ and denote this sub-matrix by $\bar{\mathbf{A}}^{(0)}$. From the second part of the proof of Theorem 4.1.1, the following relationship holds:

$$\mathbb{S}_{\bar{\mathbf{A}}^{(0)}} \cap \mathbb{R}_+^{N-R_c} = \emptyset \tag{4.5.6}$$

In addition, from Lemma 3.1.1, there exists an $(N - R_c) \times 1$ positive vector $\bar{\mathbf{v}}$ such that $\bar{\mathbf{v}} \in \mathbb{S}_{\bar{\mathbf{A}}^{(0)}}^{\perp} \cap \mathbb{R}_{++}^{N-R_c}$, satisfying

$$\bar{\mathbf{v}}^T \left(\bar{\mathbf{A}}^{(0)}\right)^T \bar{\mathbf{A}}^{(0)}\bar{\mathbf{v}} = 0 \tag{4.5.7}$$

Then, we construct an $N \times 1$ nonnegative vector $\mathbf{v}$ by letting the $\ell_i^{(0)}$-th entry be given by the i-th entry of $\bar{\mathbf{v}}$ and the other R_c entries be zero. Such vector $\mathbf{v}$ satisfies $\mathbf{v}^T\mathbf{A}^T\mathbf{A}\mathbf{v} = 0$. Therefore, if

$$\begin{aligned}\mathbf{x}_0 \in &\left\{\mathbf{x} : 0 \leq x_i \leq \frac{\tau}{\sqrt{\lambda_{\max}}}, 1 \leq i \leq N, \xi > 0\right\} \\ &\subseteq \left\{\mathbf{x} : 0 \leq \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2, \mathbf{x} \in \mathbb{R}_+^N\right\}\end{aligned} \tag{4.5.8}$$

then,

$$(\mathbf{x}_0 + \xi\mathbf{v})^T \mathbf{A}^T\mathbf{A}(\mathbf{x}_0 + \xi\mathbf{v}) = \mathbf{x}_0^T\mathbf{A}^T\mathbf{A}\mathbf{x}_0 \leq \tau^2 \tag{4.5.9}$$

holds for any positive numbers ξ and τ. Therefore,

$$(\mathbf{x}_0 + \xi\mathbf{v}) \in \left\{\mathbf{x} : 0 \leq \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2, \mathbf{x} \in \mathbb{R}_+^N\right\} \tag{4.5.10}$$

This completes the proof of Statement 1).

Proof of Statement 2): Again, from the second part of the proof of Theorem 4.1.1, we have

$$\mathbb{S}_{\bar{\mathbf{A}}^{(0)}} \cap \mathbb{R}_+^{N-R_c} = \emptyset \tag{4.5.11}$$

Then, Theorem 4.1.1 indicates that $\bar{\mathbf{A}}^{(0)}$ is zero-cover. This completes the proof of Statement 2) as well as Theorem 4.5.1. □

From Theorem 4.5.1, it can be easily seen that the following property is true.

Property 4.5.1. *If an $M \times N$ real matrix $\mathbf{A}$ is zero-cover, then, there exists an $N \times 1$*

vector $\mathbf{v}$ *with all entries being positive such that*

$$\left\{\mathbf{x}+\xi\mathbf{v} : 0 \leq x_i \leq \frac{\tau}{\sqrt{\lambda_{\max}}}, 1 \leq i \leq N, \xi > 0\right\}$$
$$\subseteq \left\{\mathbf{x} : 0 \leq \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2, \mathbf{x} \in \mathbb{R}_+^N\right\}$$

where τ *is a constant, and* $\lambda_{\max}$ *is the maximum eigenvalue of* $\mathbf{A}^T\mathbf{A}$. ■

4.6 Conclusion

In this chapter, we have presented some matrix properties related to cover theory. These can be used to determine the cover order. We also proposed the échelon transformation of the matrix. Based on the specific échelon form of the matrix, an efficient and effective method has been developed to determine the cover order for any given matrix. The structure of zero-cover matrix has been investigated, which can be useful to obtain the feasible solution for the system of linear equations with nonnegativity constraints on solutions.

Furthermore, we also develop the concepts about non-negatively linear independence and non-negatively linear dependence, with which we could gain a deeper insight on the system of linear equations with nonnegativity constraints on solutions and the meaning of the cover order.

Additionally, using the concept of generalized inverse of the matrix, we have discovered a dual relationship such that a column vector $\mathbf{a}_i$ of $\mathbf{A}$ can be determined as covered or not by examining the associated column vector in a related matrix $\tilde{\mathbf{A}}$.

Finally, the properties of inner hyper-rectangle have been revealed and are found to be closely related to zero-cover.

Chapter 5

Hyper-Rectangle Cover Theory and Linear Programming Problem

Linear programming (LP, also known as linear optimization) is a method for obtaining optimal results (such as maximum profit or lowest cost) in a mathematical model whose requirements are represented by linear relationships. Formally, a linear program is an optimization problem whose objective function is linear in variables, subject to linear equality and linear inequality constraints. Geometrically speaking, the feasible region determined by the constraints of a linear program is a convex polytope, which is a set defined as the intersection of finite half spaces and each half space is defined by a linear inequality. The real-valued linear objective function is defined on this polytope. The method for solving a linear programming problem is to seek a point in this polytope where the objective function has the smallest (or largest) value if there exist such a point. The application of linear programming can be found in a wide range of fields of study. It is widely used in mathematics, in business and economics, as well as in some engineering problems. Practically, there are a number of industries that use linear programming models for optimization, including transportation, energy, manufacturing and telecommunications. It has been shown to be useful when modeling diverse types of planning, routing, scheduling, and assignment problems.

In this chapter, we will apply the hyper-rectangle cover theory to the analysis of linear programming problem. Given a linear program, we propose a novel method based on the results we obtained in Chapter 3 to determine the necessary and sufficient condition to guarantee the non-empty feasibility set of the problem. In addition, we transform the linear program into a new structure with which we perform the échelon transformation to the augmented matrix. With the échelon form of the new augmented matrix and the condition to guarantee the non-negative solutions to the system of linear equations, the three possible consequences for any given linear program: 1) The linear program is infeasible; 2) The linear program is feasible but (objective) unbounded; 3) The linear program is feasible and has an optimal solution, are discussed in detail. On the other hand, we can also determine the feasibility and the boundedness of the linear program by verifying the cover order of this augmented matrix. Then, with the specific properties of zero-cover matrix, we are able to derive a series of feasible solutions to the linear programming problem.

Furthermore, we compare our proposed method with the widely used simplex method in solving linear programming problem and apply the cover method in solving the Klee-Minty cube problem which shows that the worst-case complexity of simplex method is exponential.

5.1 Linear Programming Problem

In this section, we will present a systematic procedure using the concept of hyper-rectangle cover for solving linear programming problems with non-negativity constraints on the variables.

5.1.1 Linear Program Form with Non-negativity Constraints

A linear program is an optimization problem where all involved functions, both in the objective and in the constraints, are linear in the variable $\mathbf{x}$ and it is one of the most used mathematical techniques in today's modern applications.

Consider the linear programming problem with non-negativity constraints on the variables:

Problem 5.1.1.

$$
\begin{aligned}
\max \quad & c_1x_1 + c_2x_2 + \cdots + c_nx_n \\
\text{subject to} \quad & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1N}x_N \leq b_1 \\
& a_{21}x_1 + a_{22}x_2 + \cdots + a_{2N}x_N \leq b_2 \\
& \vdots \\
& a_{(M-1)1}x_1 + a_{(M-1)2}x_2 + \cdots + a_{(M-1)N}x_N \leq b_{(M-1)} \\
& a_{M1}x_1 + a_{M2}x_2 + \cdots + a_{MN}x_N \leq b_M \\
& x_1, x_2, \cdots, x_N \geq 0
\end{aligned}
$$

The objective function is a linear combination of variables $x_1, x_2, \cdots, x_N$ and the coefficients are $c_1, c_2, \cdots, c_N$ respectively. Let $\mathbf{c} = (c_1, c_2, \cdots, c_N)^T$ and $\mathbf{x} =$

$(x_1, x_2, \cdots, x_N)^T$. Then the objective function becomes $\mathbf{c}^T\mathbf{x}$.

In a matrix-vector notation, the linear program with non-negative constraints on the variables becomes:

Problem 5.1.2 (LP Problem)**.**

$$\begin{aligned} \max \quad & \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

where $\mathbf{A} \in \mathbb{R}^{M\times N}$ and $\mathbf{b} \in \mathbb{R}^M$.

In a linear program, the objective function can either be minimization or maximization, while its constraints may include any combination of linear inequalities and equalities. There are several standard forms available in the literature, which may provide different advantages depending upon the circumstance. Here we will use the following standard form:

Problem 5.1.3 (A Standard Form of LP Problem)**.**

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

where $\mathbf{A}$ is a $M \times N$ real matrix, with $M < N$ and $\mathbf{b} \in \mathbb{R}^M$. The requirement $M < N$ ensures that in general there is an infinite number of solutions to the linear equation $\mathbf{A}\mathbf{x} = \mathbf{b}$, which leaves the degrees of freedom for non-negative and optimal solutions.

The variable vector $\mathbf{x}$ and the coefficient vector $\mathbf{c}$ are real column vectors with N elements. The objective function of $\mathbf{x}$ to be optimized is a linear combination of x_i in $\mathbf{x}$.

We note that the above linear programming problem has a specific form, which satisfies the following two conditions:

1. All variables x_i in $\mathbf{x}$ are constrained to be non-negative.

2. All constraints, except for the nonnegativity of the variables, are in the form of equalities.

As we know, any linear programming problem can be transformed to the form in Problem 5.1.3, which usually requires adding extra variables and constraints. By adding a so-called slack variable, the inequality can be transformed into an equivalent equality. The slack variables essentially take up the slack in the inequalities. Thus, the following results will be valid for all linear programs.

We will assume that $\mathbf{A}$ has full rank since redundant or inconsistent linear equations can always be detected and removed if so desired, thus, the rows of $\mathbf{A}$ are linearly independent which ensures that the equations in $\mathbf{Ax} = \mathbf{b}$ are consistent for any right-hand side column vector $\mathbf{b}$.

The feasibility set of above linear program is:

$$\mathcal{F}_1 = \{\mathbf{x} \in \mathbb{R}^N : \mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}\} \subset \mathbb{R}^N \tag{5.1.1}$$

A linear program is infeasible if its feasibility set is empty, otherwise, it is feasible. Those points that fall within this feasibility set are referred to as feasible points, from which we attempt to search for the optimal solution $\mathbf{x}^*$ that minimizes the objective function $\mathbf{c}^T\mathbf{x}$.

5.1.2 Hyper-Rectangle Cover Theory and the Feasibility Set of Linear Programming Problem

According to the necessary and sufficient condition for the existence of non-negative solution for the non-homogeneous linear equations system that we arrived in Theorem 3.2.1 in Chapter 3, we can easily obtain the following result which can be used to verify the emptiness of the feasibility set of a linear program:

Theorem 5.1.1. *Given a linear programming problem:*

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

where $\mathbf{A} \in \mathbb{R}^{M\times N}$*, with* $M < N$ *and* $\mathbf{b} \in \mathbb{R}^M$*. The feasibility set* $\mathcal{F}_1$ *of above linear programming problem is nonempty if and only if* $R_c(\tilde{\mathbf{A}}) \leq R_c(\mathbf{A})$*, where* $\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{A} & -\mathbf{b} \end{pmatrix}$. ■

Let us denote the objective function $\mathbf{c}^T\mathbf{x}$ as z and it is treated as a constant in the later discussion. By adding the objective function into the constraints, Problem 5.1.3 can be transformed into:

Problem 5.1.4.

$$\begin{array}{rc} \min & z \\ \text{subject to} & \begin{pmatrix} \mathbf{A} \\ \mathbf{c}^T \end{pmatrix} \mathbf{x} = \begin{pmatrix} \mathbf{b} \\ z \end{pmatrix} \\ & \mathbf{x} \geq \mathbf{0} \end{array}$$

To simplify our discussion, we use $\mathbf{A}_c$ and $\mathbf{b}(z)$ to denote the coefficient matrix of $\mathbf{x}$ and the right hand side column vector in the equation of above LP problem, i.e.,

$$\mathbf{A}_c = \begin{pmatrix} \mathbf{A} \\ \mathbf{c}^T \end{pmatrix}, \quad \mathbf{b}(z) = \begin{pmatrix} \mathbf{b} \\ z \end{pmatrix} \tag{5.1.2}$$

Multiplying the right hand side of the equation in Problem 5.1.4 by (-1) and padding the product vector to the right of $\mathbf{A}_c$, we form the matrix:

$$\mathbf{A}(z) = \begin{pmatrix} \mathbf{A} & -\mathbf{b} \\ \mathbf{c}^T & -z \end{pmatrix} \tag{5.1.3}$$

Here, $\mathbf{A}(z)$ is a $(M+1) \times (N+1)$ real matrix.

The feasibility set of above linear program is:

$$\mathcal{F}_2 = \{\mathbf{x} \in \mathbb{R}^N : \mathbf{A}_c\mathbf{x} = \mathbf{b}(z),\ \mathbf{x} \geq \mathbf{0}\} \subset \mathbb{R}^N \tag{5.1.4}$$

By applying the échelon transformation that we proposed in Chapter 4 to the augmented matrix $\mathbf{A}(z)$ without changing the position of its last row and last column, we have:

$$\mathbf{A}(z) \rightarrow \begin{pmatrix} \mathbf{I} & \mathbf{B} & \mathbf{f}z + \mathbf{g} \end{pmatrix} \tag{5.1.5}$$

where $\mathbf{I}$ is an $(M+1) \times (M+1)$ identity matrix, $\mathbf{B}$ is an $(M+1) \times (N-M-1)$ real matrix, $\mathbf{f}$ and $\mathbf{g}$ are both column vectors with $M+1$ elements.

In the following, for the simplicity of analysis, we separate $\mathbf{A}(z)$ into two parts and let the left-hand side part be:

$$\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{I}_{(M+1)\times(M+1)} & \mathbf{B}_{(M+1)\times(N-M-1)} \end{pmatrix}, \tag{5.1.6}$$

$\tilde{\mathbf{A}}$ is an $(M+1) \times N$ real matrix and let the last column of the échelon form of $\mathbf{A}(z)$ be:

$$\tilde{\mathbf{b}} = \mathbf{f}z + \mathbf{g}, \tag{5.1.7}$$

where $\tilde{\mathbf{b}} \in \mathbb{R}^{M+1}$ and each element in $\tilde{\mathbf{b}}$ is:

$$\tilde{b}_i = f_i z + g_i \tag{5.1.8}$$

which is in the form of a linear function of z for $i = 1,\ 2,\ \cdots,\ M+1$.

Property 5.1.1. *From Theorem 3.2.1, in order to have an non-empty feasibility set for the linear programming problem 5.1.4, adding* $\tilde{\mathbf{b}}$ *to the right hand side of* $\tilde{\mathbf{A}}$ *does not increase the cover order of* $\tilde{\mathbf{A}}$*. In other words, the cover order of* $\mathbf{A}(z)$ *is less than or equal to the cover order of* $\mathbf{A}_c$*.* ■

A linear program is unbounded if the feasibility set of it is not empty but its objective function can be made arbitrarily "well-behaved". More specifically, if a linear program is a minimization problem and unbounded, then its objective value can be made arbitrarily small while maintaining feasibility. In other words, the objective function can be negative infinity within the feasibility set.

Similarly, for an unbounded maximization problem, the objective value can be positive infinity within the feasibility set.

For unbounded objective function over the feasiblility set, the following property holds:

Property 5.1.2. *In a minimization problem, if there exists an 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.* ■

Similarily, we can also obtain the result for maximization problem.

Property 5.1.3. *In a maximization problem, if there exists an uncovered variable has a positive 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 Linear Programming Problem Solution

Based on Property 5.1.1, we will analyze the possibilities of the solutions and the optimal value of the objective function of the linear programming problem from the following three cases: 1) $\tilde{\mathbf{A}}$ has full-cover; 2) $0 < R_c(\tilde{\mathbf{A}}) < N$; 3)$\tilde{\mathbf{A}}$ has zero-cover. The detailed analysis as follows.

Theorem 5.2.1. *If $\tilde{\mathbf{A}}$ has full-cover and the matrix $\mathbf{B}$ in $\tilde{\mathbf{A}}$ is a non-negative matrix, then the linear programming problem has optimal solution if and only if $\tilde{b}_i = f_i z + g_i \leq 0$, for $i = 1,\ 2,\ \cdots,\ R_r$. By solving these inequalities, we will have the range of z, which is:*

$$\max\{-\frac{g_i}{f_i^{(+)}}, i \in \mathcal{I}\} \leq z \leq \min\{-\frac{g_i}{f_i^{(-)}}, i \in \mathcal{I}\}, \tag{5.2.1}$$

where $f_i^{(+)}$ and $f_i^{(-)}$ are the positive and negative terms in $\tilde{\mathbf{b}}_{\mathcal{I}}$ respectively. ■

Proof. The proof of above theorem follows directly from Property 5.1.1. □

It should be noted that if the constraint of z in Theorem 5.2.1 is contradictory, i.e., if

$$\min\{-\frac{g_i}{f_i^{(-)}}, i \in \mathcal{I}\} < \max\{-\frac{g_i}{f_i^{(+)}}, i \in \mathcal{I}\}, \tag{5.2.2}$$

then the feasibility set of this linear program is empty:

$$\begin{aligned} \mathcal{F}_2 &= \{\mathbf{x} \in \mathbb{R}^N : \mathbf{A}_c\mathbf{x} = \mathbf{b}(z),\ \mathbf{x} \geq \mathbf{0}\} \\ &= \emptyset \end{aligned} \tag{5.2.3}$$

In other words, we are not able to find any feasible solution to this linear program in this case.

If there is no lower bound of z, i.e., $\max\{-\frac{g_i}{f_i^{(+)}}, i \in \mathcal{I}\}$ in Eq. (5.2.1) can be negative infinity, then the objective function in this minimization problem will be 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:

$$\max z = \min\{-\frac{g_i}{f_i^{(-)}}, i \in \mathcal{I}\}. \tag{5.2.4}$$

If it does not exist, then the corresponding maximization problem is unbounded and if the constraint is contradictory, the feasible domain is empty which is same with the minimization case.

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 section of Chapter 4, the first non-negative row vector of $\tilde{\mathbf{A}}$ 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 $f_i z + g_i \leq 0$, for $i \in \mathcal{I}$ and a candidate minimal value of z is:

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

Step 2 If z_0 satisfies:

$$\max\{-\frac{g_k}{f_k^{(-)}}, k \in \{1, 2, \cdots, R_r\} \backslash \mathcal{I}\} \leq z_0 \leq \min\{-\frac{g_k}{f_k^{(+)}}, k \in \{1, 2, \cdots, R_r\} \backslash \mathcal{I}\}$$

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

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

Otherwise, there exists some $k \in \{1, 2, \cdots, R_r\} \backslash \mathcal{I}$ 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., the variable to introduce into the basis) by:

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

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

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

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 $\mathbf{B}$ is a non-negative matrix in the new échelon form, then the process ends and the optimal value is obtained, which is

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

Otherwise, the process continue.

Step 7 Go to step 1.

■

The whole pivot process in 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 will become $\mathbf{e}_{\bar{k}}$. Next, exchanging the position of the $(M + j_k)$-th column and the $\bar{k}$-th column. After this process, $f_{\bar{k}}z + g_{\bar{k}}$ will be negative and the structure of identity matrix ahead is reserved.

The above computational procedures of cover theory in solving LP problem are summarized in the following flow diagram- Figure 5.1.

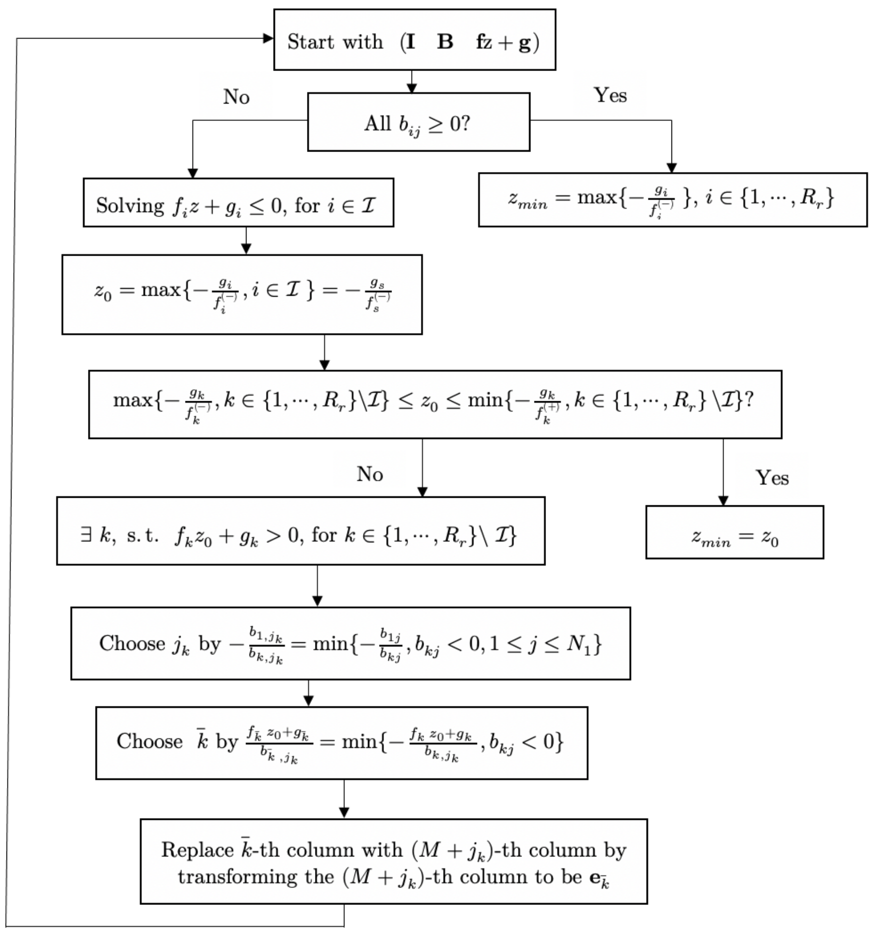

Figure 5.1: Flow diagram of the cover method

We can approach the LP problem by performing the above steps in cover method and we are making an attempt to show that such algorithm will not go to an infinite loop by improving the lower bound of z_0 in each Step 1 and it will go to the optimal value for stop. However, we have not succeed yet.

For a better understanding of the above procedures, let us consider the following example where $\tilde{\mathbf{A}}$ has full cover but the matrix $\mathbf{B}$ in $\tilde{\mathbf{A}}$ is not a non-negative matrix.

Example 5.2.1.

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

let:

$$z = \mathbf{c}^T\mathbf{x} = -x_1 - x_2$$

By adding the subject function into the constraints, we 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 $\tilde{\mathbf{A}}(z)$*, we have*

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

*Since we exchange the position of the first two rows during this transformation, the corresponding position of variables are also exchanged. According to Theorem**??**, by investigating* $f_i z + g_i \leq 0$*, for* $i = 1$ *and* $i = 3$*, we have*

$$-9 - z \leq 0$$
$$-21 - 3z \leq 0$$

By solving the above two inequalities, we will have a candidate optimal value of z*, which is*

$$z_0 = \max\{-9, -7\} = -7$$

since $z_0 \leq \min\{-\frac{g_2}{f_2}\} = -\frac{9}{2}$*. As a result, the optimal value of the objective function is* $z^* = -7$*, and the corresponding optimal solution is* $\mathbf{x}^* = \begin{pmatrix} 5 & 2 & 0 & 0 \end{pmatrix}^T$ ■

Similarly, for the case when $0 < R_c(\mathbf{A}) < N$, 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 first s rows of $\tilde{\mathbf{A}}$, where s is obtained through échelon transformation. And $k \in \{1, 2, \cdots, s\} \backslash \mathcal{J}$.

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

Theorem 5.2.2. *For full rank matrix $\tilde{\mathbf{A}}$, if it has zero-cover, then the linear programming problem is feasible but unbounded.* ■

Proof. Since adding any column to the right hand side of a zero-cover matrix with full-rank still yields a matrix with zero-cover, the feasible set $\mathcal{F}_2$ is always non-empty in this case. However, according to Theorem 4.2.1, a zero-cover matrix can be transformed to the form which has at least one negative column or has one non-positive column, but the position where the zero lies will be negative in some other column of it. Then by Property 5.1.2, the objective function is unbounded over the feasible domain in this case. □

5.3 Feasible Solution of the Linear Programming Problem

Let us first review the following linear programming problem that we discussed in the last section:

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

and the augmented matrix of $\mathbf{x}$ is

$$\mathbf{A}(z) = \begin{pmatrix} \mathbf{A} & -\mathbf{b} \\ \mathbf{c}^T & -z \end{pmatrix}$$

where the matrix $\mathbf{A}$, column vector $\mathbf{b}$ and $\mathbf{c}$ are components in Problem 5.1.3. By applying the échelon transformation to $\mathbf{A}(z)$ without changing the position of the last row and the last column, we obtain:

$$\mathbf{A}(z) \rightarrow \left(\begin{array}{ccc} \mathbf{I}_{(M+1)\times(M+1)} & \mathbf{B}_{(M+1)\times(N-M-1)} & \mathbf{f}z + \mathbf{g} \end{array} \right).$$

As we discussed in the last section, we divide the matrix $\mathbf{A}(z)$ into two parts, and let:

$$\begin{aligned} \tilde{\mathbf{A}} &= \left(\begin{array}{cc} \mathbf{I}_{(M+1)\times(M+1)} & \mathbf{B}_{(M+1)\times(N-M-1)} \end{array} \right), \\ \tilde{\mathbf{b}} &= \mathbf{f}z + \mathbf{g}. \end{aligned}$$

Thus

$$\left(\begin{array}{ccc} \mathbf{I}_{(M+1)\times(M+1)} & \mathbf{B}_{(M+1)\times(N-M-1)} & \mathbf{f}z + \mathbf{g} \end{array} \right) = \left(\begin{array}{cc} \tilde{\mathbf{A}} & \tilde{\mathbf{b}} \end{array} \right)$$

is an échelon form matrix. Then by Eq. (4.3.1), the échelon form can be divided in to the following blocks:

$$\left(\begin{array}{cc} \tilde{\mathbf{A}} & \tilde{\mathbf{b}} \end{array} \right) = \begin{pmatrix} \mathbf{I}_s & \mathbf{0} & \mathbf{B}^{(1)} & \mathbf{B}^{(3)} & \tilde{\mathbf{b}}^{(1)} \\ \mathbf{0} & \mathbf{I}_{(M+1-s)} & \mathbf{B}^{(2)} & \mathbf{B}^{(4)} & \tilde{\mathbf{b}}^{(2)} \end{pmatrix} \tag{5.3.1}$$

where s is obtained through échelon transformation. Then according to Theorem 3.1.1, in the following linear equations:

$$\begin{pmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{b}} \end{pmatrix} \mathbf{x} = \mathbf{0}, \tag{5.3.2}$$

the covered variables x_i are all zeros, as a result, we can ignore those covered column vectors in $\begin{pmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{b}} \end{pmatrix}$, which correspond to $\begin{pmatrix} \mathbf{I}_s & \mathbf{B}^{(1)} \\ \mathbf{0} & \mathbf{B}^{(2)} \end{pmatrix}$. Since $\mathbf{B}^{(3)}$ and $\tilde{\mathbf{b}}^{(1)}$ in Eq. (5.3.1) are zero matrix and zero vector respectively. In the following, we only need to consider the remaining part of it, i.e., the lower side of the matrix in Eq. (5.3.1), which is

$$\begin{pmatrix} \mathbf{I}_{(M+1-s)} & \mathbf{B}^{(4)} & \tilde{\mathbf{b}}^{(2)} \end{pmatrix} \tag{5.3.3}$$

Without loss of generality, we denote this part as:

$$\begin{aligned} \begin{pmatrix} \bar{\mathbf{A}} & \bar{\mathbf{b}} \end{pmatrix} &= \begin{pmatrix} \mathbf{I} & \bar{\mathbf{B}} & \bar{\mathbf{b}} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{I}_{(M+1-s)} & \mathbf{B}^{(4)} & \tilde{\mathbf{b}}^{(2)} \end{pmatrix} \end{aligned} \tag{5.3.4}$$

The cover order of $\begin{pmatrix} \mathbf{I} & \bar{\mathbf{B}} & \bar{\mathbf{b}} \end{pmatrix}$ is zero. Thus, in order to obtain the feasible solution for the linear programming problem, we only need to solve the following system of homogeneous linear equations, where the non-negative variable $\bar{\mathbf{x}}$ is the uncovered part in $\mathbf{x}$:

$$\begin{pmatrix} \mathbf{I} & \bar{\mathbf{B}} & \bar{\mathbf{b}} \end{pmatrix} \bar{\mathbf{x}} = \mathbf{0} \tag{5.3.5}$$

For the simplicity of the discussion, we can assume the size of $\mathbf{I}$ is $m \times m$, $\bar{\mathbf{B}}$ is a $m \times (n-m)$ real matrix and $\bar{\mathbf{b}}$ is a $m \times 1$ column vector.

From Theorem 4.2.1, we know that the zero-cover matrix can be transformed to the form which contains at least one negative column vector, or has one non-positive column vector, but the position where the zero lies will be negative in some other column of it. Without loss of generality, we can assume the negative column appears in the first column of $\bar{\mathbf{B}}$, i.e.,

$$\bar{\mathbf{b}}_1^T = (\bar{b}_{11},\ \bar{b}_{21},\ \cdots,\ \bar{b}_{m1})^T \tag{5.3.6}$$

is a negative column vector, i.e., $\bar{b}_{i1} < 0$, for $i = 1,\ \cdots,\ m$.

Then the following procedure enables us to obtain a series of feasible solutions to the linear programming problem.

Suppose $\bar{\mathbf{x}} = (\bar{x}_1,\ \cdots,\ \bar{x}_m,\ \bar{x}_{m+1}, \bar{x}_{m+2},\ \cdots,\ \bar{x}_n,\ \bar{x}_{n+1})$, where the first m elements in $\bar{\mathbf{x}}$: $\bar{x}_1,\ \cdots,\ \bar{x}_m$ 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 $\begin{pmatrix} \mathbf{I} & \bar{\mathbf{B}} & \bar{\mathbf{b}} \end{pmatrix} \bar{\mathbf{x}}$.

Then, according to Eq. (5.3.5), the first m elements in $\bar{\mathbf{x}}$ can be expressed as a linear combination of the last $m - n + 1$ elements: $\bar{x}_{m+1},\ \bar{x}_{m+2},\ \cdots,\ \bar{x}_n,\ \bar{x}_{n+1}$ as follows:

$$\begin{aligned}
\bar{x}_1 &= -\bar{b}_{11}\bar{x}_{m+1} - \bar{b}_{12}\bar{x}_{m+2} - \cdots - \bar{b}_{1(n-m)}\bar{x}_n - \bar{b}_1\bar{x}_{n+1} \\
\bar{x}_2 &= -\bar{b}_{21}\bar{x}_{m+1} - \bar{b}_{22}\bar{x}_{m+2} - \cdots - \bar{b}_{2(n-m)}\bar{x}_n - \bar{b}_2\bar{x}_{n+1} \\
&\vdots \\
\bar{x}_m &= -\bar{b}_{m1}\bar{x}_{m+1} - \bar{b}_{m2}\bar{x}_{m+2} - \cdots - \bar{b}_{m(n-m)}\bar{x}_n - \bar{b}_m\bar{x}_{n+1}
\end{aligned} \tag{5.3.7}$$

In order to obtain a linearly independent feasible solution set, we first let the vector $(\bar{x}_{m+1}, \ \cdots, \ \bar{x}_n, \ \bar{x}_{n+1})^T$ 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. Also, according to Eqs. (5.3.7), in order to satisfy the nonnegativity constraints on the variable $\bar{x}_i$, $i = 1, \ \cdots, \ n+1$, we let

$$L = \max\left\{-\frac{\bar{b}_{i2}}{\bar{b}_{11}}, \ -\frac{\bar{b}_{i3}}{\bar{b}_{21},}, \ \cdots, \ -\frac{\bar{b}_{i(n-m)}}{\bar{b}_{m1}}, \ -\frac{\bar{b}_i}{\bar{b}_{i1}}\right\}, \ i = 1, 2, \cdots, m. \tag{5.3.8}$$

Then we are able to get a set of linear independent basic feasible solution $\boldsymbol{\alpha}_i$, for $i = 1, \ 2, \ \cdots, \ n-m-1, \ n-m$:

$$\begin{aligned} \boldsymbol{\alpha}_1 &= (-\bar{b}_{11}L - \bar{b}_{12}, \cdots, -\bar{b}_{m1}L - \bar{b}_{m2}, L, 1, 0, \cdots, 0)^T \\ \boldsymbol{\alpha}_2 &= (-\bar{b}_{11}L - \bar{b}_{13}, \cdots, -\bar{b}_{m1}L - \bar{b}_{m3}, L, 0, 1, \cdots, 0)^T \\ &\vdots \\ \boldsymbol{\alpha}_{n-m-1} &= (-\bar{b}_{11}L - \bar{b}_{1n}, \cdots, -\bar{b}_{m1}L - \bar{b}_{mn}, L, 0, 0, \cdots, 0, 1)^T \\ \boldsymbol{\alpha}_{n-m} &= (-\bar{b}_{11}L - \bar{b}_1, \cdots, -\bar{b}_{m1}L - \bar{b}_m, L, 0, 0, \cdots, 0, 1)^T \end{aligned} \tag{5.3.9}$$

Thus the solution of the following linear equations:

$$\begin{pmatrix} \mathbf{I} & \bar{\mathbf{B}} & \bar{\mathbf{b}} \end{pmatrix} \bar{\mathbf{x}} = \mathbf{0} \tag{5.3.10}$$

can be represented by the convex combination of this set of basic feasible solution $\boldsymbol{\alpha}_i$, where $i = 1, \ 2, \ \cdots, \ n-m$, i.e.,

$$\bar{\mathbf{x}} = k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_{n-m}\boldsymbol{\alpha}_{n-m} \tag{5.3.11}$$

with $k_1,\ k_2,\ \cdots,\ k_{n-m}$ being any positive real value, and

$$k_1 + k_2 + \cdots + k_{n-m} = 1. \tag{5.3.12}$$

By adding the covered variables to $\bar{\mathbf{x}}$, we will obtain a series of feasible solutions to Problem 5.1.4.

5.4 Comparison with Simplex Method in Solving Linear Programming Problem

5.4.1 Simplex Method

In 1947, Dantzig developed the *simplex method*, which is the first algorithm that solves the linear programming problem efficiently in most cases. Geometrically speaking, the feasible region of a linear program is a convex polytope, which is defined as the intersection of finitely many half spaces. The objective function is a real-valued affine function defined over this polytope. An extreme point or vertex of the convex polytope is known as *basic feasible solution* (BFS). The linear programming problem is to find an extreme point of this polytope where the objective function has the smallest (or largest) value if such an extreme point exists. By moving along the edge of the polytope to the extreme points, the simplex method identifies the extreme points with better objective values, i.e., it proceeds by moving from one feasible solution to another at each step, improving the value of the objective function. The process continues until the optimum objective value is reached, or an unbounded edge is visited. Due to the finite number of extreme points in the polytope, for an LP problem having a non-empty feasible region, the algorithm always terminates after a finite number of transitions and it terminates in one, and only one, of the following possible situations:

1. by determining an optimal solution;
2. by demonstrating that there is no feasible solution; or
3. by demonstrating that the objective function is unbounded over the feasible region.

Before presenting a formal and general version of the simplex method. The canonical form of linear programs is introduced. It satisfies the following:

1. All variables are constrained to be non-negative.
2. All constraints are expressed as equalities, except for the non-negative constraints of variables.
3. The righthand-side coefficients in the constraints are all non-negative.
4. The basic variable is isolated in each constraint. The variable which is isolated in a given constraint will not appear in any other constraint, and the coefficient of it in the objective function is zero.

Since any linear program can be transformed to its canonical form, the following discussion on the simplex method is valid for any general linear programs. Given a linear program in the form of Problem 5.1.1, it can be transformed into the following canonical form (8):

Problem 5.4.1 (Canonical Form)**.**

$$
\begin{aligned}
x_1 + \bar{a}_{1,m+1}x_{m+1} + \cdots + \bar{a}_{1s}x_s + \cdots + \bar{a}_{1n}x_n &= \bar{b}_1, \\
x_2 + \bar{a}_{2,m+1}x_{m+1} + \cdots + \bar{a}_{2s}x_s + \cdots + \bar{a}_{2n}x_n &= \bar{b}_2, \\
&\vdots \\
x_r + \bar{a}_{r,m+1}x_{m+1} + \cdots + \bar{a}_{rs}x_s + \cdots + \bar{a}_{rn}x_n &= \bar{b}_r, \\
&\vdots \\
x_m + \bar{a}_{m,m+1}x_{m+1} + \cdots + \bar{a}_{ms}x_s + \cdots + \bar{a}_{mn}x_n &= \bar{b}_m, \\
(-z) + \bar{c}_{m+1,m+1}x_{m+1} + \cdots + \bar{c}_s x_s + \cdots + \bar{c}_n x_n &= -\bar{z}_0, \\
x_j \geq 0 \ (j = 1, 2, \cdots, n)&
\end{aligned}
$$

The data $\bar{a}_{ij}$, $\bar{b}_i$, $\bar{z}_0$ and $\bar{c}_j$ in Problem 5.4.1 are known. They are either the original data or the data that updated during the transformation. $x_1, \ x_2, \ \cdots, \ x_m$ are assumed to be the basic variables. Moreover, since the above problem is in the canonical form, we have the righthand-side coefficients $b_i \geq 0$ for $i = 1, 2, \ \cdots, \ m$.

The following steps are the essential computation steps of the simplex method in solving linear programming problem.

Simplex Algorithm (Maximization Form)

Step 1 The problem is initially in canonical form and all $\bar{b}_i \geq 0$.

Step 2 If $\bar{c}_j \leq 0$ for $j = 1, 2, \cdots, n$, then the algorithm ends and the optimal value is obtained. Otherwise, there exists some $\bar{c}_j > 0$ and the algorithm continues.

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

$$\bar{c}_s = \max_j \{\bar{c}_j,\ \bar{c}_j > 0\}.$$

If $\bar{a}_{is} \leq 0$ for $i = 1, 2, \cdots, m$, then stop; the primal problem is unbounded.

If there exists $\bar{a}_{is} > 0$ for some $i = 1, 2, \cdots, m$, then the process continues.

Step 4 Choose row r to pivot in (i.e., the variable to drop from the basis) by the ratio test:

$$\frac{\bar{b}_r}{\bar{a}_{rs}} = \min_i \left\{ \frac{\bar{b}_i}{\bar{a}_{is}},\ \bar{a}_{is} > 0 \right\}$$

Step 5 Replace the basic variable in row r with variable s and re-establish the canonical form (i.e., pivot on the coefficient $\bar{a}_{rs}$).

Step 6 Go to step 1.

The pivot steps are summarized pictorially in Fig.5.2 and the last tableau specifies the new values for the data after the pivot.

x_1	$\cdots$	x_r	$\cdots$	x_m	x_{m+1}	$\cdots$	x_s	$\cdots$	x_n	
1					$\bar{a}_{1,m+1}$	$\cdots$	$\bar{a}_{1s}$	$\cdots$	$\bar{a}_{1n}$	$\bar{b}_1$
	$\ddots$				$\vdots$				$\vdots$	$\vdots$
		1			$\bar{a}_{r,m+1}$	$\cdots$	$\boxed{\bar{a}_{rs}}$	$\cdots$	$\bar{a}_{rn}$	$\bar{b}_r$
			$\ddots$		$\vdots$				$\vdots$	$\vdots$
				1	$\bar{a}_{m,m+1}$	$\cdots$	$\bar{a}_{ms}$	$\cdots$	$\bar{a}_{mn}$	$\bar{b}_m$
					$\bar{c}_{m+1}$	$\cdots$	$\bar{c}_s$	$\cdots$	$\bar{c}_n$	$-\bar{z}_0$

↓Normalization

1					$\bar{a}_{1,m+1}$	$\cdots$	$\bar{a}_{1s}$	$\cdots$	$\bar{a}_{1n}$	$\bar{b}_1$
	$\ddots$				$\vdots$				$\vdots$	$\vdots$
		$\left(\frac{1}{\bar{a}_{rs}}\right)$			$\left(\frac{\bar{a}_{r,m+1}}{\bar{a}_{rs}}\right)$	$\cdots$	1	$\cdots$	$\left(\frac{\bar{a}_{rn}}{\bar{a}_{rs}}\right)$	$\left(\frac{\bar{b}_r}{\bar{a}_{rs}}\right)$
			$\ddots$		$\vdots$				$\vdots$	$\vdots$
				1	$\bar{a}_{m,m+1}$	$\cdots$	$\bar{a}_{ms}$	$\cdots$	$\bar{a}_{mn}$	$\bar{b}_m$
					$\bar{c}_{m+1}$	$\cdots$	$\bar{c}_s$	$\cdots$	$\bar{c}_n$	$-\bar{z}_0$

↓ Elimination of x_s

1		$-\left(\frac{\bar{a}_{1s}}{\bar{a}_{rs}}\right)$			$\bar{a}_{1,m+1} - \bar{a}_{1s}\left(\frac{\bar{a}_{r,m+1}}{\bar{a}_{rs}}\right)$	$\cdots$	0	$\cdots$	$\bar{a}_{1n} - \bar{a}_{1s}\left(\frac{\bar{a}_{rn}}{\bar{a}_{rs}}\right)$	$\bar{b}_1 - \bar{a}_{1s}\left(\frac{\bar{b}_r}{\bar{a}_{rs}}\right)$
	$\ddots$				$\vdots$				$\vdots$	$\vdots$
		$\left(\frac{1}{\bar{a}_{rs}}\right)$			$\left(\frac{\bar{a}_{r,m+1}}{\bar{a}_{rs}}\right)$	$\cdots$	1	$\cdots$	$\left(\frac{\bar{a}_{rn}}{\bar{a}_{rs}}\right)$	$\frac{\bar{b}_r}{\bar{a}_{rs}}$
			$\ddots$		$\vdots$				$\vdots$	$\vdots$
		$-\left(\frac{\bar{a}_{ms}}{\bar{a}_{rs}}\right)$		1	$\bar{a}_{m,m+1} - \bar{a}_{ms}\left(\frac{\bar{a}_{r,m+1}}{\bar{a}_{rs}}\right)$	$\cdots$	0	$\cdots$	$\bar{a}_{mn} - \bar{a}_{ms}\left(\frac{\bar{a}_{rn}}{\bar{a}_{rs}}\right)$	$\bar{b}_m - \bar{a}_{ms}\left(\frac{\bar{b}_r}{\bar{a}_{rs}}\right)$
		$-\left(\frac{\bar{c}_s}{\bar{a}_{rs}}\right)$			$\bar{c}_{m+1} - \bar{c}_s\left(\frac{\bar{a}_{r,m+1}}{\bar{a}_{rs}}\right)$	$\cdots$	0	$\cdots$	$\bar{c}_n - \bar{c}_s\left(\frac{\bar{a}_{rn}}{\bar{a}_{rs}}\right)$	$-\bar{z}_0 - \bar{c}_s\left(\frac{\bar{b}_r}{\bar{a}_{rs}}\right)$

Figure 5.2: Algebra for a pivot operation.

It is observed that the new value for z is:

$$z^{new} = \bar{z}_0 + \bar{c}_s\left(\frac{\bar{b}_r}{\bar{a}_{rs}}\right). \tag{5.4.1}$$

By the choice of the variable x_s introduced into the basis, we have $\bar{c}_s > 0$. Since $\bar{b}_r \geq 0$ and $\bar{a}_{rs} > 0$, this implies that $z^{\text{new}} \geq z^{\text{old}}$. As a result, after each pivot, the objective function will reach a better or same value. Additionally, if $\bar{b}_r > 0$, then the objective value strictly increases after every pivot.

Finally, in a finite number of iterations, the simplex method will terminate and it will show that there is no feasible solution; finds an optimal solution; or show that the objective function is unbounded over the feasible region.

In practice, the simplex method has shown remarkable efficiency. However, in 1972, Klee and Minty gave an example, the Klee–Minty cube (35), showing that the worst-case complexity of simplex method is exponential time. From then on, almost every variation of the method has shown to give poor results for a specific class of linear programs. It is an open question if there is a variation of simplex method having polynomial computation complexity.

5.4.2 Comparison

In the simplex method, the objective value z in the canonical tableau of linear programming problem is regarded as a variable whereas in the cover method, it is treated as a constant. Given a linear program, the cover method first transforms the problem into the canonical form of Problem 5.1.4, and then $\mathbf{A}(z)$ is then 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 5.2.1. 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 the computation complexity of it.

Consider an $M \times N$ real matrix $\mathbf{A}$, where $M < N$ and $\mathbf{A}$ has full rank. The complexity of transforming $\mathbf{A}$ into an échelon form is $\mathcal{O}(M^2N)$. Then we find out all non-negative row vectors in $\mathbf{B}$ and select the one with the greatest number of non-zero elements. In the échelon transformation process, this selected row is supposed to be moved to the first row. Meanwhile, the non-zero elements in this row are expected to be moved to the forward side of $\mathbf{B}$ while performing the corresponding column permutation such that the identity matrix structure could be preserved. In the following step, without considering the columns of $\mathbf{A}$ corresponding to those non-zero elements in this row, perform the same transformation on the remainder of $\mathbf{A}$ to obtain its échelon form. However, in order to simplify the computation complexity of solving LP problem, once the non-negative row vector with the largest number of positive terms is identified, instead of performing the corresponding row and column permutations to obtain the canonical échelon form of $\mathbf{A}$, we can analyze the remaining part of $\mathbf{A}$ without considering the columns of $\mathbf{A}$ corresponding to those non-zero elements in this row directly. This can simplify the computational complexity by reducing the number of transformation operations. The complexity of this structural re-arrangement process is $\mathcal{O}(M^2(N - M))$. As a result, the total computation complexity of solving this kind of LP problem by cover method is $\mathcal{O}(M^2N)$.

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

Property 5.1.1 indeed proposes a condition for which the optimum objective value appears. The optimum BFS can also be determined by examining the row where the optimum objective value is encountered.

5.4.3 Klee-Minty Cube Example

In this section, we will apply the cover theory to solve the Klee–Minty cube problem. The problem is formulated as follows (26):

$$\begin{aligned}
\min\ -2^{N-1}x_1 - 2^{N-2}x_2 - \cdots - 2x_{N-1} - x_N \\
x_1 \le 5 \\
4x_1 + x_2 \le 25 \\
8x_1 + 4x_2 + x_3 \le 125 \\
\vdots \\
2^N x_1 + 2^{N-1}x_2 + \cdots + 4x_{N-1} + x_N \le 5^N \\
x_1, x_2, \cdots, x_N \ge 0
\end{aligned}$$

The linear program has N variables, N inequality constraints except for the nonnegativity constraints on variables x_i, for $i = 1,\ 2,\ \cdots,\ N$, and 2^N extreme points. Let us denote the objective function as z, then we have:

$$-2^{N-1}x_1 - 2^{N-2}x_2 - \cdots - 2x_{N-1} - x_N = z.$$

By introducing the slack variables $s_i \ge 0, i = 1,\ 2,\ \cdots,\ N$, the inequality constraints can be transformed into equations. In addition, padding the above z-equation into the coefficient matrix of variables x_i and s_i, for $i = 1,\ 2,\ \cdots,\ N$, we will have the following matrix:

$$\mathbf{A}(z) = \begin{pmatrix} 1 & 0 & \cdots & 0 & 2^N & 2^{N-1} & \cdots & 4 & 1 & -5^N \\ 0 & 1 & \cdots & 0 & 2^{N-1} & 2^{N-2} & \cdots & 1 & 0 & -5^{N-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 1 & 0 & \cdots & 0 & 0 & -5 \\ 0 & 0 & \cdots & 0 & -2^{N-1} & -2^{N-2} & \cdots & -2 & -1 & -z \end{pmatrix}$$

The first row of $\mathbf{A}(z)$ consists the largest number of positive elements in all row vectors in it, and

$$\arg\max_{N+1\leq i\leq 2N}\left(-\frac{A(z)_{(N+1)i}}{A(z)_{1i}}\right) = 2N$$

Thus, in the next step, let the $2N$-th column vector of $\mathbf{A}(z)$ be $\mathbf{e}_{N+1}$, where $\mathbf{e}_{N+1} \in \mathbb{R}^{N+1}$ and the elements in it are all zeros except the $N+1$-th term in it is one. Then moving the the $2N$-th column to the $N+1$-th column of the resulted matrix, we then have the échelon form of $\mathbf{A}(z)$:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & 2^N - 2^{N-1} & 2^{N-1} - 2^{N-2} & \cdots & 2 & -5^N - z \\ 0 & 1 & \cdots & 0 & 0 & 2^{N-1} & 2^{N-2} & \cdots & 1 & -5^{N-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 1 & 0 & \cdots & 0 & -5 \\ 0 & 0 & \cdots & 0 & 1 & 2^{N-1} & 2^{N-2} & \cdots & 2 & z \end{pmatrix}$$

Perform the same permutation to the variables in $\mathbf{x}$ and according to Theorem 5.2.1, the optimal value of the objective function is

$$z^* = -5^N$$

and the corresponding optimal solution is

$$\mathbf{x}^* = (0, 0, \cdots, 0, 5^N)^T$$

It is observed that if the simplex method formulated by Dantzig is applied in solving this Klee–Minty cube problem, and suppose the initial vertex for the simplex method is the origin, then the method goes through all 2^N extreme points (55), finally reaching the optimal vertex $(0,\ 0,\ \cdots,\ 0,\ 5^N)^T$. Thus, showing that the worst-case complexity of simplex method is exponential time.

In the following, we provide a detailed pivot sequence of simplex method for $N = 3$ (26), which goes through all 8 extreme points, starting at the origin, to show the exponential time complexity of simplex method in solving the Klee-Minty cube problem. Let $s_1,\ s_2,\ s_3$ be the slack variables.

Initial Tableau

Basic	Nonbasic			RHS
	x_1	x_2	x_3	
s_1	1^*			5
s_2	4	1		25
s_3	8	4	1	125
$-z$	4	2	1	0

Tableau 1

Basic	Nonbasic			RHS
	s_1	x_2	x_3	
x_1	1			5
s_2	-4	1^*		5
s_3	-8	4	1	85
$-z$	-4	2	1	-20

Tableau 2

Basic	Nonbasic			RHS
	x_1	x_2	x_3	
x_1	1^*			5
x_2	-4	1		5
s_3	8	-4	1	65
$-z$	4	-2	1	-30

Tableau 3

Basic	Nonbasic			RHS
	x_1	s_2	x_3	
s_1	1			5
x_2	4	1		25
s_3	-8	-4	1^*	25
$-z$	-4	-2	1	-50

Tableau 4

Basic	Nonbasic			RHS
	x_1	s_2	s_3	
s_1	1^*			5
x_2	4	1		25
x_3	-8	-4	1	25
$-z$	4	2	-1	-75

Tableau 5

Basic	Nonbasic			RHS
	s_1	s_2	s_3	
x_1	1			5
x_2	-4	1^*		5
x_3	8	-4	1	65
$-z$	-4	2	-1	-95

Tableau 6

Basic	Nonbasic			RHS
	s_1	x_2	s_3	
x_1	1^*			5
s_2	-4	1		5
x_3	-8	4	1	85
$-z$	4	-2	-1	-105

Tableau 7

Basic	Nonbasic			RHS
	x_1	s_2	x_3	
s_1	1^*			5
s_2	4	1		25
x_3	8	4	1	125
$-z$	-4	-2	-1	-125

Figure 5.3: Klee-Minty polytope shows exponential time complexity of simplex method

5.5 Conclusion

On the basis of the échelon form and the corresponding results on the system of linear equations with nonnegativity constraints on solutions, we can verify whether or not the feasibility set of the linear programming problem is empty. Then the solutions and the optimal values of the linear programs have the following possibilities, (i) it has optimal bounded solution, (ii) it is feasible but unbounded, or (iii) it has infinite unbounded optimal solution. These correspond to the following scenarios: full cover, zero cover or the cover order in between, and have been analyzed in detail. Moreover, with the échelon form and the specific structure of zero-cove matrix, a series of feasible solutions of any given linear programming problem can be obtained. We also include a comparison of cover method and simplex method in solving the linear programming problem in this chapter. We apply the cover method to solve the Klee-Minty cube problem with great efficiency.

Chapter 6

Cover Length

In Chapter 4, we have derived a method to determine the cover order of any given real matrix $\mathbf{A}$. In this chapter, we will propose a method to obtain the cover length c_i of the covered variables x_i in $\mathbf{Ax}$, where $\mathbf{A} \in \mathbb{R}^{M \times N}$ and $\mathbf{x} \in \mathbb{R}_+^N$. In addition, we also find that there is a strong relationship between the cover length problem and the non-negative least square problem, i.e., the NNLS problem can be reconstructed as a problem of determining the cover length of the corresponding variable. Therefore, it is possible to obtain an analytical result of the NNLS problem by applying the method to determine the cover length. We will also include a discussion of various algorithms for solving the NNLS problem and the method developed in this section.

6.1 Cover Length Determination of the Covered Variable

We first encountered the concept of *cover length* in Definition 2.1.1. In this section, we propose a method which can be used to determine the cover length of the covered variable x_i in $\mathbf{Ax}$.

In general, given an $M \times N$ real matrix $\mathbf{A}$, and real column vector $\mathbf{x}$ with N non-negative elements in it, if the i-th variable x_i is covered, for $i = 1,\ 2,\ \cdots,\ N$, then the cover length c_i of it can be obtained by solving the following optimization problem:

Problem 6.1.1. *Let* $\mathbf{A}$ *be an* $M \times N$ *real matrix,* $\mathbf{x} = \{x_1, x_2,\ \cdots,\ x_N\}^T \in \mathbb{R}_+^N$ *and* x_N *be covered in* $\mathbf{Ax}$*.*

$$\begin{aligned} \max \quad & x_N \\ \text{subject to} \quad & \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq 1 \end{aligned} \tag{6.1.1}$$

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

It is obvious that if the maximum value of x_N does not exist, then the value of x_N can be very large under the constraints meaning that x_N is not covered within the feasible domain. Otherwise, if Problem 6.1.1 has an optimal value, i.e., the maximum value of x_N exists within the feasibility region, then x_N is covered and according to the definition of cover length of the covered variable in Definition 2.1.1, the maximum value of x_N which satisfies the inequality constraints: $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq 1$, is the cover length c_N of the variable of x_N.

In order to solve the above optimization problem, let us form a Lagrangian function corresponding to the constrained optimization problem as follows:

$$L(\mathbf{x}, \boldsymbol{\lambda}) = -x_N - \sum_{n=1}^{N} \lambda_n x_n + \frac{\lambda_{N+1}}{2}(\mathbf{x}^T \mathbf{A}^T \mathbf{A}\mathbf{x} - 1), \tag{6.1.2}$$

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 of Problem 6.1.1 is that the following Karush-Kuhn-Tucker(KKT) conditions (54) must be satisfied:

$$\begin{aligned}
\nabla L(\mathbf{x}, \boldsymbol{\lambda})|_{\mathbf{x}=\mathbf{x}^*, \boldsymbol{\lambda}=\boldsymbol{\lambda}^*} &= -\mathbf{e}_N - \boldsymbol{\lambda}^* + \lambda^*_{N+1}\mathbf{A}^T\mathbf{A}\mathbf{x}^* = \mathbf{0}, \\
x^*_n \lambda^*_n &= 0 \qquad \text{for } n = 1, 2, \cdots, N, \\
\lambda^*_{N+1}\big((\mathbf{x}^*)^T\mathbf{A}^T\mathbf{A}\mathbf{x}^* - 1\big) &= 0, \\
(\mathbf{x}^*)^T\mathbf{A}^T\mathbf{A}\mathbf{x}^* &\leq 1, \\
\mathbf{x}^* &\geq \mathbf{0}, \\
\lambda^*_{N+1} &\geq 0, \\
\boldsymbol{\lambda}^* &\geq \mathbf{0},
\end{aligned} \tag{6.1.3}$$

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 note that $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x}$ can be rewritten as follows,

$$\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = p_{NN}\Big(x_N + \frac{\bar{\mathbf{p}}^T_N\bar{\mathbf{x}}_N}{p_{NN}}\Big)^2 + \bar{\mathbf{x}}^T_N\Big(\bar{\mathbf{P}}_{NN} - \frac{\bar{\mathbf{p}}_N\bar{\mathbf{p}}^T_N}{p_{NN}}\Big)\bar{\mathbf{x}}_N \tag{6.1.4}$$

where $\mathbf{P} = \mathbf{A}^T\mathbf{A}$ is a $N \times N$ 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 N-th row and N-th column from it, $\bar{\mathbf{p}}_N$ is the $(N-1)\times 1$ vector generated by deleting N-th entry from the N-th row of $\mathbf{P}$ and $\bar{\mathbf{x}}_N$ denotes the $(N-1)\times 1$ non-negative vector obtained by deleting the N-th entry from $\mathbf{x}$.

Therefore, we can have another way to represent the KKT conditions of the optimization problem as follows:

$$\begin{aligned}
-\bar{\boldsymbol{\lambda}}_N^* + \lambda_{N+1}^*\Big(\big(x_N^* + \frac{\bar{\mathbf{p}}_N^T\bar{\mathbf{x}}_N^*}{p_{NN}}\big)\bar{\mathbf{p}}_N + \big(\mathbf{P}_{NN} - \frac{\bar{\mathbf{p}}_N\bar{\mathbf{p}}_N^T}{p_{NN}}\big)\bar{\mathbf{x}}_N^*\Big) &= \mathbf{0}\\
-1 - \lambda_N^* + \lambda_{N+1}^*\Big(p_{NN}x_N^* + \bar{\mathbf{p}}_N^T\bar{\mathbf{x}}_N^*\Big) &= 0\\
x_n^*\lambda_n^* &= 0\\
\lambda_{N+1}^*\big((\mathbf{x}^*)^T\mathbf{A}^T\mathbf{A}\mathbf{x}^* - 1\big) &= 0\\
(\mathbf{x}^*)^T\mathbf{A}^T\mathbf{A}\mathbf{x}^* &\leq 1\\
\mathbf{x}^* &\geq \mathbf{0}\\
\lambda_{N+1}^* &\geq 0\\
\boldsymbol{\lambda}^* &\geq \mathbf{0},
\end{aligned} \tag{6.1.5}$$

for $n = 1, 2, \cdots, N$. Here, $\bar{\boldsymbol{\lambda}}_N^*$ is the $(N-1)\times 1$ vector generated by deleting N-th entry from $\boldsymbol{\lambda}^*$ and $\bar{\mathbf{x}}_N^*$ denotes the $(N-1)\times 1$ vector generated by deleting N-th entry from $\mathbf{x}^*$. Since $x_N^* \neq 0$, then by $x_n^*\lambda_n^* = 0$, for $n = 1, 2, \cdots, N$, we have $\lambda_N^* = 0$. In addition, in order to satisfy the second equation in the above KKT conditions, we have $\lambda_{N+1}^* \neq 0$ and thus we have:

$$x_N^* = \lambda_{N+1}^*. \tag{6.1.6}$$

Using the KKT conditions in Eqs.(6.1.5), the solution to Problem 6.1.1 is given in the following theorem:

Theorem 6.1.1. *Let* $\mathbf{A}$ *be an* $M \times N$ *real matrix with its rank being* R_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 nn-th element* $[\mathbf{A}^T\mathbf{A}]_{nn}$*, such that the following two conditions are satisfied simultaneously:*

1. $\mathbf{P}^{-1}_{i_j=n;i_1 i_2 \cdots i_r}\mathbf{e}_{i_j=n} \geq \mathbf{0}$ *and* $[\mathbf{P}^{-1}_{i_j=n;i_1 i_2 \cdots i_r}\mathbf{e}_{i_j=n}]_{i_j=n} > 0$;
2. $\det(\mathbf{P}_{i_j=n\rightarrow k;i_1 i_2 \cdots i_r}) \geq 0$ *for* $k = 1, 2, \cdots, N$ *but* $k \neq i_1, i_2, \cdots, i_r$*, where* $\mathbf{P}_{i_j=n\rightarrow k;i_1 i_2 \cdots i_r}$ *denotes an* $r \times r$ *sub-matrix of* $\mathbf{A}^T\mathbf{A}$ *maintaining all the entries of* $\mathbf{P}_{i_j=n;i_1 i_2 \cdots i_r}$ *but having* $(p_{k,i_1}, p_{k,i_2}, \cdots, p_{k,i_r})$ *as a new row replacing the old row:* $(p_{i_j=n,i_1}, p_{i_j=n,i_2}, \cdots, p_{i_j=n,i_r})$.
3. $c_n = \sqrt{\left[\left(\mathbf{P}_{i_j=n;i_1 i_2 \cdots i_r}\right)^{-1}\right]_{nn}}$

■

Proof. We denote the PSD matrix $\mathbf{A}^T\mathbf{A}$ as $\mathbf{P}$ in the following discussion. The KKT conditions of Problem 6.1.1 can be simplified as follows:

$$\begin{aligned} \mathbf{Px} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0} \\ \mathbf{b} &\geq \mathbf{0}, b_N > 0 \\ x_i b_i &= 0 \text{ for } i = 1, 2, \ \cdots, \ N-1 \end{aligned} \tag{6.1.7}$$

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 $\mathbf{x}$. Then we are able to find a $|\bar{\mathcal{N}}| \times |\bar{\mathcal{N}}|$ sub-matrix $\bar{\mathbf{P}}$ of $\mathbf{P}$, such that the following equality holds:

$$\bar{\mathbf{P}}\bar{\mathbf{x}} = \mathbf{e}_{|\bar{\mathcal{N}}|} \tag{6.1.8}$$

where all the elements x_i in $\bar{\mathbf{x}}$ are 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 lose of generality, we can let

$$\mathbf{T} = \{\mathbf{t}_1, \; \cdots, \; \mathbf{t}_{r-1}, \bar{\mathbf{P}}_{|\bar{\mathcal{N}}|}\}, \tag{6.1.9}$$

and we will have

$$\mathbf{T}\tilde{\mathbf{x}} = \mathbf{e}_{|\bar{\mathcal{N}}|}, \tag{6.1.10}$$

where $\tilde{\mathbf{x}} \geq \mathbf{0}$, $\tilde{\mathbf{x}} \in \mathbb{R}^r$ and the last element of $\tilde{\mathbf{x}}$: $\tilde{x}_r > 0$. This can be proved in the following:

1. Let $\mathbf{T}$ be the smallest set which containing the $|\bar{\mathcal{N}}|$-th column of $\bar{\mathbf{P}}$: $\bar{\mathbf{P}}_{|\bar{\mathcal{N}}|}$, s.t., $\mathbf{e}_{|\bar{\mathcal{N}}|} \in \text{cone}\, \mathbf{T}$, where

$$\begin{aligned} \text{cone}\, \mathbf{T} &= \text{cone}\, \{\mathbf{t}_1, \cdots, \mathbf{t}_{\mathrm{r}-1}, \bar{\mathbf{P}}_{|\bar{\mathcal{N}}|}\} \qquad (6.1.11) \\ &= \{\theta_1 \mathbf{t}_1 + \cdots + \theta_{r-1}\mathbf{t}_{r-1} + \theta_r \bar{\mathbf{P}}_{|\bar{\mathcal{N}}|} | \theta_i \geq 0 \text{ for i} = 1, \cdots, \mathrm{r}\} \end{aligned}$$

2. $\mathbf{T}$ is linearly independent, otherwise there are real value μ_j, such that,

$$\sum_{j=1}^{r-1} \mu_j t_j + \mu_r \bar{\mathbf{P}}_{|\bar{\mathcal{N}}|} = 0. \tag{6.1.12}$$

And there are $\lambda_j \geq 0$, $\lambda_r > 0$, such that,

$$\sum_{j=1}^{r-1} \lambda_j t_j + \lambda_r \bar{\mathbf{P}}_{|\bar{\mathcal{N}}|} = \mathbf{e}_{|\bar{\mathcal{N}}|}, \tag{6.1.13}$$

By multiplying the Eq.(6.1.12) with α on both side and adding the product into Eq.(6.1.13) , we can obtain the following equality:

$$\sum_{j=1}^{r-1} (\mu_j + \alpha\lambda_j) t_j + (\mu_r + \alpha\lambda_r) \bar{\mathbf{P}}_{|\bar{\mathcal{N}}|} = \mathbf{e}_{|\bar{\mathcal{N}}|}. \tag{6.1.14}$$

If $\mu_r \geq 0$, then let

$$\alpha = \max_{1 \leq j \leq r-1} \left\{ -\frac{\lambda_j}{\mu_j}, \mu_j > 0 \right\} = -\frac{\lambda_i}{\mu_i}. \tag{6.1.15}$$

Thus for every $1 \leq j \leq r-1$, we have:

$$\lambda_j + \alpha\mu_j \geq 0, \tag{6.1.16}$$

while

$$\lambda_i + \alpha\mu_i = 0. \tag{6.1.17}$$

Then we can have a new $\tilde{\mathbf{x}} \in \mathbb{R}^{r-1}$ with the $r-1$-the element being positive while others being all non-negative:

$$\tilde{x}_{r-1} = \lambda_r + \alpha\mu_r > 0, \quad (6.1.18a)$$

$$\tilde{x}_k = \lambda_j + \alpha\mu_j \geq 0, \; for \;\; 1 \leq j \leq r-1, and \; j \neq i \quad (6.1.18b)$$

where $\tilde{x}_{r-1}$ is the last element in new $\tilde{\mathbf{x}}$, λ_r, μ_r, λ_j, μ_j are elements in Eq.(6.1.14) and $\tilde{x}_k$ is the k-th element in the new $\tilde{\mathbf{x}}$, where $1 \leq k \leq r-2$.

If $\mu_r < 0$, then let

$$\alpha = \max_{1 \leq j \leq r-1} \left\{ -\frac{\lambda_j}{\mu_j}, \mu_j > 0 \right\} = -\frac{\lambda_i}{\mu_i}. \quad (6.1.19)$$

Then we can have a new $\tilde{\mathbf{x}} \in \mathbb{R}^{r-1}$ with the $r-1$-the element being positive while others being all non-negative in the same manner as the case when $\mu_r \geq 0$.

As a consequence, we can always find a smaller set $\widetilde{\mathbf{T}}$ which contains $\bar{\mathbf{P}}_{|\bar{\mathcal{N}}|}$ such that $\mathbf{e}_{|\bar{\mathcal{N}}|} \in$ cone $\widetilde{\mathbf{T}}$, which contradicts to the assumption that $\mathbf{T}$ is the smallest set that containing $\bar{\mathbf{P}}_{|\bar{\mathcal{N}}|}$ such that $\mathbf{e}_{|\bar{\mathcal{N}}|} \in$ cone $\mathbf{T}$.

As a result, $\mathbf{T}$ is linearly independent. According to the constraints of $\mathbf{x}$, $\tilde{\mathbf{x}}$ should be equivalent to $\bar{\mathbf{x}}$ and $\mathbf{T} = \bar{\mathbf{P}}$. Thus $\bar{\mathbf{P}}$ is invertible. Since we have $\bar{\mathbf{P}}\bar{\mathbf{x}} = \mathbf{e}_{|\bar{\mathcal{N}}|}$, where $\bar{\mathbf{x}} \geq \mathbf{0}, \bar{\mathbf{x}} \in \mathbb{R}^r$ and $\bar{x}_r > 0$, then we will have

$$\mathbf{P}^{-1}_{i_j=n;i_1 i_2 \cdots i_r} \mathbf{e}_{i_j=n} \geq \mathbf{0}, \quad (6.1.20)$$

and the $(i_j = n)$-th element in it is positive.

Till now, the first statement has been proved.

By using the new row $(p_{k,i_1}, p_{k,i_2}, \;\cdots\;, \; p_{k,i_r})$ to replace the old row $(p_{i_j=n,i_1}, p_{i_j=n,i_2},$

$\cdots$, $p_{i_j=n,i_r}$) in $\mathbf{P}_{i_1 i_2 \cdots i_r}$, we will have:

$$\mathbf{P}_{i_j=n\to k;i_1 i_2 \cdots i_r}\bar{\mathbf{x}} = b_k \mathbf{e}_{i_j=n} \tag{6.1.21}$$

To simplify the expression, we denote $\mathbf{P}_{i_j=n\to k;i_1 i_2 \cdots i_r}$ as $\bar{\mathbf{P}}_k$. If $\bar{\mathbf{P}}_k$ is inverible, then by Cramer's Rule (36), we can get:

$$x_n = \frac{b_k \left|\bar{\mathbf{P}}_{(r-1)\times(r-1)}\right|}{\left|\bar{\mathbf{P}}_k\right|} \tag{6.1.22}$$

where $\bar{\mathbf{P}}_{(r-1)\times(r-1)}$ is the $(r-1)$-th order leading principle sub-matrix of the matrix $\mathbf{P}_{i_1 i_2 \cdots i_r}$. Since $\mathbf{P}_{i_1 i_2 \cdots i_r}$ is a positive definite matrix, and for PD matrix, its leading principal minors are all positive (6). The i-th leading principal minor of a matrix $\mathbf{P}_{i_1 i_2 \cdots i_r}$ is the determinant of its upper-left $i \times i$ sub-matrix. According to the property of PD matrix, we can get $\det(\bar{\mathbf{P}}_{(r-1)\times(r-1)}) > 0$. Since $b_k \geq 0$, $x_n > 0$, then from Eq.(6.1.22), we can obtain that $\det(\bar{\mathbf{P}}_k) > 0$. And 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 $\sqrt{\left[\left(\mathbf{P}_{i_j=n;i_1 i_2 \cdots i_r}\right)^{-1}\right]_{nn}}$. □

From the above discussion, it is observed that Theorem 6.1.1 can also be applied to determine whether the i-th variable x_i in $\mathbf{x}$ is covered or not. In other words, we can also conclude that if we are not able to find out a principal sub-matrix that can satisfy the conditions listed in Theorem 6.1.1, then there is no optimal value to Problem 6.1.1. Thus we say that the corresponding variable is not covered over the feasible region.

In the following, We will examine an example to illustrate the features of the above method to obtain the cover length of a covered variable before presenting more details.

Example 6.1.1. *Given the folllowing* 4×4 *matrix* $\mathbf{A}$

$$\mathbf{A} = \begin{pmatrix} -3 & -2 & -5 & -2 \\ 3 & -5 & 0 & -4 \\ 1 & 3 & 1 & -3 \\ 2 & 2 & 1 & 4 \end{pmatrix}$$

we can obtain its corresponding PSD matrix

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

In order to determine the cover length of the covered variable x_4*, we find the principal sub-matrix of* $\mathbf{P}$ *which can satisfy all the conditions listed in Theorem 6.1.1. We will start our examination from checking all the* 2×2 *principle sub-matrices. Specifically, we only need to consider the* 2×2 *principal sub-matrices which contain negative* *element in the right upper side corner, since only this kind of* 2×2 *principal sub-matrices may satisfy the condition that the last column of the inverse of it is a non-negative vector and the second element in it is positive. As a result, for the second-order principal sub-matrices we only need to verify the following principal sub-matrix:*

$$\mathbf{P}_{14} = \begin{pmatrix} 23 & -1 \\ -1 & 45 \end{pmatrix}.$$

1. *We verify that* $\mathbf{P}_{14}$ *is invertible and the last column of the its inverse is a positive column vector.*

2. *Then we replace the second row in* $\mathbf{P}_{14}$ *with the other rows resulting in the following two sub-matrices:*

$$\mathbf{P}_{4\to2;14} = \begin{pmatrix} 23 & -1 \\ -2 & 23 \end{pmatrix}, \mathbf{P}_{4\to3;14} = \begin{pmatrix} 23 & -1 \\ 18 & 11 \end{pmatrix}.$$

   Then we check the determinant of them and they are verified to be both non-negative.

From above discussion, we can see that the invertible 2×2 *principal sub-matrix* $\mathbf{P}_{14}$ *satisfies all the conditions in Theorem 6.1.1 and we have:*

$$\mathbf{P}_{14}^{-1} = \begin{pmatrix} \frac{45}{1034} & \frac{1}{1034} \\ \frac{1}{1034} & \frac{23}{1034} \end{pmatrix}.$$

As a result, we can conclude that the cover length of x_4 is

$$c_4 = \sqrt{\frac{23}{1034}}.$$

■

Lemma 6.1.1. *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 $c_n = \frac{1}{\sqrt{[\mathbf{A}^T\mathbf{A}]_{nn}}}$.*
2. *If $\mathbf{A}^T\mathbf{A}$ has full rank and all the entries in the n-th column of $\left(\mathbf{A}^T\mathbf{A}\right)^{-1}$ are positive, then, we have $c_n = \sqrt{\left[(\mathbf{A}^T\mathbf{A})^{-1}\right]_{nn}}$.*

■

Proof. To prove the first statement, we consider an $M \times N$ real matrix $\mathbf{A}$ and a non-negative real vector $\mathbf{x}$ with N elements in it. We can rewrite $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x}$ as follows:

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

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 in Statement 1), i.e., all the entries in $\mathbf{A}^T\mathbf{A}$ are positive, then we can always have that $\bar{\mathbf{x}}^T\bar{\mathbf{A}}^T\bar{\mathbf{A}}\bar{\mathbf{x}} \geq 0$, $\bar{\mathbf{x}}^T\bar{\mathbf{A}}^T\mathbf{a}_n x_n \geq 0$, $\mathbf{a}_n^T\bar{\mathbf{A}}\bar{\mathbf{x}}x_n \geq 0$ and $\mathbf{a}_n^T\mathbf{a}_n x_n^2 \geq 0$. Thus for any given positive real-valued number $\tau > 0$, letting $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \leq \tau^2$ allows us to arrive at $\mathbf{a}_n^T\mathbf{a}_n x_n^2 \leq \tau^2$, which gives us $x_n \leq \frac{\tau}{\sqrt{\mathbf{a}_n^T\mathbf{a}_n}}$. This observation tells us that for any $\mathbf{x} \in \mathbb{R}_+^N$ and a given matrix $\mathbf{A} \in \mathbb{R}^{M\times N}$, of which all the entries in its PSD matrix $\mathbf{A}^T\mathbf{A}$ are positive, satisfying $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \le \tau^2$, the maximum achievable value of x_n is $\frac{1}{\sqrt{[\mathbf{A}^T\mathbf{A}]_{nn}}}$. Therefore, in this case, according to the definition of cover length in Definition 2.1.1, we have the cover length of the n-th variale x_n is given by $c_n = \frac{1}{\sqrt{[\mathbf{A}^T\mathbf{A}]_{nn}}}$.

The second statement can be obtained from Theorem 6.1.1 directly. □

6.2 Further Discussion on Cover Length Determination

Property 6.2.1. *Suppose there are k negative elements, $P_{i_1N}, P_{i_2N}, \cdots, P_{i_kN}$ in the N-th column of $\mathbf{P}$. Then we only need to examine whether there exist $\mathbf{P}_{i_jN}, j = 1, \cdots, k$ that can satisfy the conditions in Theorem 6.1.1 when checking two-by-two principal sub-matrix of $\mathbf{P}$.* ■

Furthermore, if there does not exist any two-by-two principal sub-matrix of $\mathbf{P}$ that can satisfy the conditions in Theorem 6.1.1, when we come to the checking of 3×3 principal sub-matrix of $\mathbf{P}$, we only need to consider those 3×3 principal sub-matrices of $\mathbf{P}$ whose first two elements in the third column of it contains at least one negative entry.

Property 6.2.2. *If there exists an invertible r-order principal sub-matrix $\mathbf{P}_{i_1i_2\cdots i_{r-1}N}$ of $\mathbf{P}$ including $\mathbf{P}_{NN}$ and*

$$\mathbf{P}^{-1}_{i_1i_2\cdots i_{r-1}N}\mathbf{e}_N > \mathbf{0},$$

and there exists k such that

$$\det(\mathbf{P}_{N\to k;i_1i_2\cdots N}) < 0$$

for $k = 1, 2, \cdots, N-1$ while $k \neq i_1, i_2, \cdots, i_{r-1}$, then, let $\{j_1, j_2, \cdots, j_{r-1}\}$ be an index set which is an ordered arrangement of $r-1$ distinct elements from $\{i_1, i_2, \cdots, i_{r-1}\}$. Now, if

$$\det(\mathbf{P}_{j_1j_2\cdots j_{r-2}i_{r-1}\to kj_{r-1};j_1j_2\cdots j_{r-2}i_{r-1}\to kN}) < 0$$

for all possible arrangements, then $\mathbf{P}_{i_1i_2\cdots i_{r-1}kN}$ is the $(r+1)$-order principal sub-matrix to be considered in the next step. ■

The above property provides us with some information for the determination of the next $(r+1)$ order principal sub-matrix that we need to examine based on the current situation that the present r order principal sub-matrix cannot satisfy all the conditions listed in Theorem 6.1.1. This provides us with a way to reduce the size of the algorithm for cover length determination. The following example is used to illustrate the above two properties.

Example 6.2.1. *Given a 4×4 matrix $\mathbf{A}$ as follows:*

$$\mathbf{A} = \begin{pmatrix} -2 & 3 & 5 & 4 \\ 4 & 0 & -1 & -2 \\ 0 & 0 & -2 & 3 \\ 4 & -4 & 2 & 4 \end{pmatrix}$$

and its PSD matrix is

$$\mathbf{P} = \mathbf{A}^{\mathbf{T}}\mathbf{A} = \begin{pmatrix} 36 & -22 & -6 & 0 \\ -22 & 25 & 7 & -4 \\ -6 & 7 & 34 & 24 \\ 0 & -4 & 24 & 45 \end{pmatrix}$$

We want to determine the cover length of x_4 *in* $\mathbf{Ax}$. *According to Property 6.2.1, the only* 2×2 *principal sub-matrix that we need to examine is*

$$\mathbf{P}_{24} = \begin{pmatrix} 25 & -4 \\ -4 & 45 \end{pmatrix}$$

It is invertible and $\mathbf{P}_{24}^{-1}\mathbf{e}_2 > \mathbf{0}$. *Since*

$$\det(\mathbf{P}_{21;24}) = \det \begin{pmatrix} 25 & -4 \\ -22 & 0 \end{pmatrix} < 0,$$

but

$$\det(\mathbf{P}_{23;24}) = \det \begin{pmatrix} 25 & -4 \\ 7 & 24 \end{pmatrix} > 0$$

Then we consider $\det(\mathbf{P}_{12;14}) = \det \begin{pmatrix} 36 & 0 \\ -22 & -4 \end{pmatrix}$, *which is smaller than 0. Thus according to Property 6.2.2,* $\mathbf{P}_{124}$ *is the* 3×3 *matrix that we are going to examine* *in the next step. Consider*

$$\mathbf{P}_{124} = \begin{pmatrix} 36 & -22 & 0 \\ -22 & 25 & -4 \\ 0 & -4 & 45 \end{pmatrix}$$

we have:

1. $\det(\mathbf{P}_{124}) > 0$, $\mathbf{P}_{124}^{-1}\mathbf{e}_3 > \mathbf{0}$;

2. $\det(\mathbf{P}_{123;124}) = \det \begin{pmatrix} 36 & -22 & 0 \\ -22 & 25 & -4 \\ -6 & 7 & 24 \end{pmatrix} > 0$

Thus the cover length of x_4 is

$$c_4 = \sqrt{[\mathbf{P}_{124}^{-1}]_{33}} = \frac{1}{\sqrt{13/567}}$$

■

6.3 The Cover Length Determination Problem and the NNLS Problem

The non-negative least squares (NNLS) problem is a constrained least squares regression problem in which all the variables can only have non-negative values. Specifically, the NNLS problem can be formulated as follows (39):

Problem 6.3.1 (Nonnegative Least Squares(NNLS))**.** *Given a matrix* $\mathbf{B} \in \mathbb{R}^{M \times N}$ *and a column vector* $\mathbf{b} \in \mathbb{R}^M$*, find a nonnegative vector* $\mathbf{u} \in \mathbb{R}_+^N$ *such that*

$$\begin{aligned} \min \quad & \| \mathbf{B}\mathbf{u} - \mathbf{b} \|_2^2 \\ \text{subject to} \quad & \mathbf{u} \geq \mathbf{0} \end{aligned} \tag{6.3.1}$$

■

In other words, the goal of NNLS problem is to find a nonnegative vector $\mathbf{u} \in \mathbb{R}_+^N$ which can minimize the objective function $f(\mathbf{u}) = \| \mathbf{B}\mathbf{u} - \mathbf{b} \|_2^2$. In the following, we establish a connection between the cover length determination problem and the NNLS problem, that is, by introducing a new variable, we show that the NNLS problem can be reconstructed as a problem of determining the cover length of the corresponding variable. This provides us with a method to arrive at the closed-form optimal value of the objective function of the NNLS problem.

Let

$$\tau^2 = \| \mathbf{B}\mathbf{u} - \mathbf{b} \|_2^2 \tag{6.3.2}$$

When $\tau = 0$, Problem 6.3.1 is equivalent to the problem of finding solutions for the non-homogeneous linear equations $\mathbf{B}\mathbf{u} = \mathbf{b}$ with nonnegative constraints on solutions. Let us consider the case when $\tau > 0$: By dividing τ^2 on both sides of Eq.(6.3.2), we have

$$\| \mathbf{B}\frac{\mathbf{u}}{\tau} - \mathbf{b}\frac{1}{\tau} \|_2^2 = 1. \tag{6.3.3}$$

Introducing a new variable

$$\mathbf{x} = \begin{pmatrix} \frac{\mathbf{u}}{\tau} & \frac{1}{\tau} \end{pmatrix}^T. \tag{6.3.4}$$

The origin NNLS problem can be transformed into:

Problem 6.3.2.

$$\begin{aligned} \max \quad & x_{N+1} \\ \text{subject to} \quad & \| \mathbf{Ax} \|_2^2 = 1 \end{aligned} \tag{6.3.5}$$

where $x_i \geq 0$ *for* $i = 1, 2, \cdots, N$, $x_{N+1} > 0$ *and* $\mathbf{A} = \begin{pmatrix} \mathbf{B} & -\mathbf{b} \end{pmatrix}$. ■

We observe that Problem 6.3.2 is consistent with Problem 6.3.1. As a result, both 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 NNLS problem. If, in case, we are not able to find the cover length of this variable, i.e., x_{N+1} is uncovered in $\mathbf{Ax}$, then we conclude that the original NNLS problem has no optimum solution.

The following example is used to illustrate the equivalence between the NNLS problem and the cover length determination problem.

Example 6.3.1. *The cover length determination problem in Example 6.2.1 is consistent with the NNLS problem:* $\min_{\mathbf{u}\in\mathbb{R}^3_+} \| \mathbf{Bu} - \mathbf{b} \|_2^2$*, where*

$$\mathbf{B} = \begin{pmatrix} -2 & 3 & 5 \\ 4 & 0 & -1 \\ 0 & 0 & -2 \\ 4 & -4 & 2 \end{pmatrix}$$

and

$$\mathbf{b} = \begin{pmatrix} -4 & 2 & -3 & -4 \end{pmatrix}^T.$$

Let

$$\begin{aligned} \tau^2 &= \| \mathbf{Bu} - \mathbf{b} \|_2^2 \\ \mathbf{x} &= \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \end{pmatrix}^T = \begin{pmatrix} \frac{\mathbf{u}}{\tau} & \frac{1}{\tau} \end{pmatrix}^T. \end{aligned}$$

The cover length of x_4 *is*

$$c_4 = \frac{1}{\sqrt{13/567}} = \frac{1}{\tau},$$

thus the optimal value of this NNLS problem is

$$\tau^2 = \| \mathbf{Bu} - \mathbf{b} \|_2^2 = (\frac{1}{c_4})^2 = 13/567.$$

■

The above example shows us how to convert the cover length determination of a desired variable into finding the optimal value of the corresponding NNLS problem and verifies the equivalence of the two problems. Moreover, on the basis of the relationship between the cover length determination problem and the NNLS problem, we are able to obtain the analytical optimal value of the NNLS problem directly for some certain types of matrices. In the following, we will use M-matrix as an example to demonstrate this conclusion. Let us first introduce the definitions and related properties of the Z-matrix and the M-matrix (46; 59; 41; 45; 5):

Definition 6.3.1 (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* $a_{ij} \le 0 \ \forall\ i \neq j, 1 \le i, j \le N$*, is a real Z-matrix.* ■

Definition 6.3.2 (M-matrix)**.** *Let* $\mathbf{A}$ *be a* $N \times N$ *real Z-matrix. That is,* $\mathbf{A} = (a_{ij})$ *where* $a_{ij} \le 0$ *for all* $i \neq j$*,* $1 \le i, j \le N$*. Then matrix* $\mathbf{A}$ *is also an M-matrix if it can be expressed in the form*

$$\mathbf{A} = s\mathbf{I} - \mathbf{T}, \tag{6.3.6}$$

where $\mathbf{T} = (t_{ij})$ *with* $t_{ij} \ge 0$*, for all* $i \neq j$*,* $1 \le i, j \le N$*, and* s *is at least as large as the maximum of the moduli of the eigenvalues of the matrix* $\mathbf{T}$*, and* $\mathbf{I}$ *is an identity matrix.* ■

Many statements that are equivalent to the definition of a non-singular M-matrix are known, and any one of these statements can serve as a definition of a non-singular M-matrix. In the following, we only mention the characterizations that we will need in our discussion.

Lemma 6.3.1. *Let* $\mathbf{A} \in \mathbb{R}^{N\times N}$ *be a Z-matrix, then the following statements are equivalent to* $\mathbf{A}$ *being a non-singular M-matrix:*

1. *All the principal minors of* $\mathbf{A}$ *are positive. That is, the determinant of each submatrix 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 all the elements in* $\mathbf{A}^{-1}$ *are all non-negative.*

■

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

Theorem 6.3.1. *Let matrix* $\mathbf{B} \in \mathbb{R}^{N\times(N-1)}$ *and vector* $\mathbf{b} \in \mathbb{R}^{N}$*. Denote* $\mathbf{A}$ *as* $\begin{pmatrix}\mathbf{B} & -\mathbf{b}\end{pmatrix}$*. Suppose* $\mathbf{A}$ *is a non-singular M-matrix, then the optimal value of the NNLS problem* $\min\limits_{\mathbf{u}\in\mathbb{R}_+^{N-1}} \| \mathbf{Bu} - \mathbf{b} \|_2^2$ *is equal to* $\frac{1}{[(\mathbf{A}^T\mathbf{A})^{-1}]_{NN}}$. ■

Proof. By assumption, $\mathbf{A}$ is a non-singular M-matrix, therefore $\mathbf{A}^T\mathbf{A}$ is invertible and all the elements in $(\mathbf{A}^T\mathbf{A})^{-1}$ are positive according to Theorem 6.3.1. When the original NNLS problem is formulated into the problem of cover length determination of x_N by applying Lemma 6.1.1, the cover length of the corresponding variable x_N is given by:

$$c_N = \sqrt{\left[(\mathbf{A}^T\mathbf{A})^{-1}\right]_{NN}}.$$

Thus, the optimal value is

$$\tau^2 = \| \mathbf{Bu} - \mathbf{b} \|_2^2 = \left(\frac{1}{c_N}\right)^2 = \frac{1}{[(\mathbf{A}^T\mathbf{A})^{-1}]_{NN}}$$

□

In mathematics, a stochastic matrix is a square matrix used to describe the transitions of a Markov chain. Each of its entries is a nonnegative real number representing a probability (25; 37; 60). It is also called a probability matrix, transition matrix, substitution matrix, or Markov matrix.

Definition 6.3.3 (Stochastic Matrix)**.** *A stochastic matrix is a non-negative square matrix, with each row summing to 1.* ■

In the following, a simple example is given as an application of Lemma 6.1.1 in obtaining the optimal value of the NNLS problem when $\mathbf{A}$ is a stochastic matrix. It shows finding the minimum value of the objective is a one-step process.

Example 6.3.2. *Let matrix* $\mathbf{B} \in \mathbb{R}^{N\times(N-1)}$ *and vector* $\mathbf{b} \in \mathbb{R}^N$. *Denote* $\mathbf{A}$ *as* $\begin{pmatrix} \mathbf{B} & -\mathbf{b} \end{pmatrix}$. *Suppose* $\mathbf{A}$ *is a stochastic matrix, then the optimal value of the NNLS problem* $\min_{\mathbf{u}\in\mathbb{R}_+^{N-1}} \| \mathbf{Bu} - \mathbf{b} \|_2^2$ *is* $\left[\mathbf{A}^T\mathbf{A}\right]_{NN}$. ■

6.4 Comparison with the Active-Set Method in Solving NNLS Problem

The first widely used algorithm for solving NNLS problems is an active-set method published by Lawson and Hanson in their 1974 book (39). There are several features of this normally used active set method. A typical example of an active-set method is implemented by the algorithm `lsqnonneg` in Matlab, aiming at obtaining an active set and arriving at an approximate solution to NNLS problem with this active set. The detailed algorithm is given as follows.

Algorithm 1 `lsqnonneg`

Input: $\mathbf{B} \in \mathbb{R}^{M\times N}$, $\mathbf{b} \in \mathbb{R}^{M}$
Output: $\mathbf{u}^* \geq \mathbf{0}$ such that $\mathbf{u}^* = \arg\max \parallel \mathbf{Bu} - \mathbf{b} \parallel_2^2$.

Initialize: $P = \emptyset$, $R = \{1, 2, \cdots, n\}$, $\mathbf{u} = \mathbf{0}$, $\mathbf{w} = \mathbf{B}^T(\mathbf{b} - \mathbf{Bu})$
repeat
Proceed if $R \neq \emptyset \wedge [\max_{i\in R}(w_i) > tolerance]$
$j = \arg\max_{i\in R}(w_i)$
Include the index j in P and remove it from R
$\mathbf{s}^P = [(\mathbf{B}^P)^T\mathbf{B}^P]^{-1}(\mathbf{B}^P)^T\mathbf{b}$
if $\min(\mathbf{s}^P) \leq 0$ **then**
  $\alpha = -\min_{i\in P}[\frac{u_i}{(u_i - s_i)}]$
  $\mathbf{u} := \mathbf{u} + \alpha(\mathbf{s} - \mathbf{u})$
  Update R and P
  $\mathbf{s}^P = [(\mathbf{B}^P)^T\mathbf{B}^P]^{-1}(\mathbf{B}^P)^T\mathbf{b}$
  $\mathbf{s}^R = \mathbf{0}$
end if
$\mathbf{u} = \mathbf{s}$
$\mathbf{w} = \mathbf{B}^T(\mathbf{b} - \mathbf{Bu})$

It 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 of $\mathbf{w}$ occurs and move this index from the inactive set to the active set. By solving the corresponding least squares problem with current active set, one non-negative solution candidate can be obtained. The active set and inactive set can be updated with current candidate solution and continue the whole process until all the elements in $\mathbf{w}$ are non-positive or the inactive set is empty.

As Lawson and Hanson show, 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 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 computational complexity is typically $\mathcal{O}(MN^2)$ per iteration (7).

Compared with active-set method, the cover length determination method takes finite operations and once we find one principal sub-matrix that can satisfy the conditions in Theorem 6.1.1, then the algorithm can stop. 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 `lsqnonneg` solution depends on a prescribed tolerance ϵ, the cover-length method yields the exact value of the function. And it performs better than `lsqnonneg` for small size matrices.

In the following, we present an example to illustrate the performance of the cover length method in solving a general NNLS problem.

Example 6.4.1. *The average running time (seconds) and average error of the `lsqnonneg` and cover length methods for the matrices and vectors randomly generated by Matlab's `rand` function are shown in the Table 6.1.*

The results listed here are averaged over 100 random samples with varying number *of columns (from 1 to 3) of* $\mathbf{B}$ *in NNLS problem. The default termination tolerance on the solution of* `lsqnonneg` *is* $10 \times \sum_{ij} |b_{ij}| \times N \times eps$*, where* $eps = 2.22 \times 10^{-16}$*,* N *is the row number of the matrix* $\mathbf{B}$*, and* b_{ij} *is the element in* $\mathbf{B}$*. Table 6.1 also includes the computation complexity (number of maximum operations) of the cover length method in solving a general NNLS problem.*

It is clear from the table that the advantage of the cover length method over `lsqnonneg` *is in the accuracy of the solution since cover length yields closed-form solutions.*

Table 6.1: Comparison between Matlab's `lsqnonneg` and cover length method in solving NNLS problem (All tests are performed on a 2.3 GHz Intel Core i5, with a memory of 8GB 2133 MHz LPDDR3.)

column number of $\mathbf{B}$		1	2	3
complexity	cover length method	6	16	589
running time (sec)	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}

■

6.5 Conclusion

In this chapter, a method is proposed to obtain the cover length of the covered variable x_i in $\mathbf{Ax}$. We also find that there is an interesting connection between the cover length determination problem and the NNLS problems. More specifically, given any NNLS problem, it can be converted to the cover length determination problem. In the process of determine the corresponding cover length of the covered variable, we are able to obtain the analytical optimal value of the NNLS problem. The importance of the application of the cover theory in the analysis of NNLS problem lies in that it solves the NNLS problem directly by investigating the matrix itself. Furthermore, for certain kinds of matrices, such as the M-matrix, the closed-form optimal value can be obtained in a more straightforward way. We also include a comparison of our proposed method and the commonly used active-set method to solve the NNLS problem so that we may have a better understanding of the advantage of cover length determination method.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

Linear system of equations with nonnegativity constraints on variables, linear programming problems and NNLS problems arise often in science, engineering and business. Nonnegativity constraints on solutions, or approximate solutions to these problems are commonly found in modern applications. For the purpose of preserving the characteristics of solutions with respect to the real-life data like images, text, and audio spectra, it is important to pay attention to nonnegativity so as to avoid spurious results. Thus, we are motivated to propose an effective technique to deal with the typical problems with nonnegativity constraints that arise in the linear algebra and optimization from the viewpoint of matrix. In this thesis, the hyper-rectangle cover theory is introduced and developed. It provides us with a new perspective of analyzing the important characteristics of nonnegativity constraints. In the process of exploring the properties of the significant concepts of cover order and cover length in the cover theory, the novel approaches and viewpoints towards these typical problems are identified.

More specifically, in the process of determining whether the variable is covered or not, the necessary and sufficient condition under which a system of homogeneous linear equations with nonnegativity constraints has a unique solution has been identified, which exactly corresponds to full-cover. As a matter of fact, we could naturally obtain the necessary and sufficient conditions for the existence of non-zero solutions for this system. In addition, with regard to the system of non-homogeneous linear equations with nonnegativity constraints on the variables, the normally used methods for analyzing the existence of the solution are mainly concerned with some other associated problems, i.e., the original problem has a solution in the required domain if and only if its associated problem has a solution. With the help of cover theory, it is possible to obtain the necessary and sufficient condition of the existence of the non-negative solutions directly by comparing the cover order of the original coefficient matrix and the augmented matrix in our form of the linear equations system. We also propose the condition for the guarantee of the unique solution, thereby filling the lacking of the characterization of uniqueness.

In the process of investigating the properties of cover order for any given matrix, we discover the equivalence between the cover order and the largest number of the positive terms of the non-negative vectors in the row space of the matrix. In addition, since performing row transformation will not alter the cover order of the matrix, we establish a specific échelon form of the matrix by carrying out a series of elementary row operations and column permutations. With this échelon transformation, we could determine the cover order for any given matrix. Specifically, for zero-cover matrices, some matrices of special form, and low-rank matrices, the échelon form of them possess particular structures. Moreover, when we examine the properties of zero-cover matrices, we gain a deeper understanding in the analysis of linear equations system with non-negative constraints on solutions. Based on the specific échelon form of matrix established, the concepts of non-negatively linear independence and non-negatively linear dependence have been developed, providing us with a deeper insight of the linear equations system with nonnegativity constraints on solutions and the meaning of cover order. We also discover the dual relationship between cover and uncover based on the related properties of the generalized inverse of matrix, i.e., the dual relationship shows that if the cover condition of vectors in the matrix is not obvious, it may be clearer when the dual matrix is inspected.

On the basis of the échelon form and the related results on the system of linear equations with nonnegativity constraints on solutions, we are able to determine whether the feasibility set of the linear programming problem is empty or not. Simplex method is widely used as a computational approach in solving linear programming problem and the procedure involves moving from one feasible solution to another. As the process progresses, the value of the objective function improves. For any given linear program, there are three possibilities: 1) the linear program is infeasible; 2) the linear program is feasible but (objective) unbounded; 3) the linear program is feasible and has an optimal solution. By employing cover theory, we obtain the conditions under which these three possibilities occur, as well as the optimal value of the objective function for each. Besides, with the property of zero-cover matrix, a series of feasible solutions to the linear programming problems can be obtained. A comparison between our proposed method and the simplex method is presented. The Klee-Minty cube problem shows that the worst-case complexity of simplex method is exponential time while our proposed method can solve it efficiently and effectively.

An analytical method is proposed to determine the cover length of the covered variable by formulating the cover length determination problem into the corresponding optimization problem. It is observed that the NNLS problem can be re-formulated into the cover length determination problem, thus a new approach is proposed to solve the NNLS problem and obtain the analytical optimal value of it. The most significant aspect of this method lies in its dependence on the structure of the matrix itself. Specifically, we only need to investigate the structure of those principal sub-matrices. For some matrices, such as the M-matrix, the optimal value could be obtained more directly. The work here also includes a comparison of various algorithms in solving NNLS problems.

7.2 Future Work

The work included in this thesis just scratches the surface of the cover theory and there are still much work to be conducted in the future to further enrich this theory.

Mathematically, we can apply the cover theory in the analysis of a great number of mathematical problems with nonnegativity constraints on the variables. By investigating the specific form of the matrix with full cover and the matrix with zero cover, we may further apply the cover theory in the analysis of non-negative matrix factorization (NMF).

From the application aspect, it is promising to apply the cover theory in the unique identification in signal processing, (especially, in the optical wireless communications), and machine learning. In particular, when we consider the large data set, the cover theory may play an important role in the reduction of the dimension.

(Specifically, there may exist two types of reduction: the first kind is that we only focus on those bounded variables — thus throwing out the uncovered variables reduces the complexity of the problem; the second kind lies in the linear equations system with nonnegativity constraints on variables — since those covered variables are zero, we can only focus on uncovered ones). This may further be applied to data recovery where the data is constrained to be non-negative.

Bibliography

[1] W. A. Adkins and S. H. Weintraub, *Algebra: an approach via module theory*, vol. 136, Springer Science & Business Media, 2012.

[2] H. Anton, *Elementary linear algebra 9 th edition.*, (1987).

[3] S. Bellavia, M. Macconi, and B. Morini, *An interior point newton-like method for non-negative least-squares problems with degenerate solution*, Numerical Linear Algebra with Applications, 13 (2006), pp. 825–846.

[4] A. Ben-Israel and T. N. Greville, *Generalized inverses: theory and applications*, vol. 15, Springer Science & Business Media, 2003.

[5] A. Berman and R. Plemmons, *Nonnegative matrices in the mathematical sciences, comput*, Sci. Appl. Math., Academic Press, NewYork, (1979).

[6] R. Bhatia, *Positive definite matrices*, Princeton university press, 2009.

[7] Å. Björck, *Numerical methods for least squares problems*, SIAM, 1996.

[8] S. P. Bradley, A. C. Hax, and T. L. Magnanti, *Applied mathematical programming*, Addison-Wesley, 1977.

[9] R. Bro and S. De Jong, *A fast non-negativity-constrained least squares algorithm*, Journal of Chemometrics: A Journal of the Chemometrics Society, 11 (1997), pp. 393–401.

[10] A. M. Bruckstein, M. Elad, and M. Zibulevsky, *On the uniqueness of nonnegative sparse solutions to underdetermined systems of equations*, IEEE Transactions on Information Theory, 54 (2008), pp. 4813–4820.

[11] S. L. Campbell and C. D. Meyer, *Generalized inverses of linear transformations*, SIAM, 2009.

[12] J. Cantarella and M. Piatek, *Tsnnls: A solver for large sparse least squares problems with non-negative variables*, arXiv preprint cs/0408029, (2004).

[13] D. Chen and R. J. Plemmons, *Nonnegativity constraints in numerical analysis*, in The birth of numerical analysis, World Scientific, 2010, pp. 109–139.

[14] M. Chu and R. Plemmons, *Nonnegative matrix factorization and applications*, Bulletin of the International Linear Algebra Society, 34 (2005), p. 26.

[15] G. Dantzig, *Linear programming and extensions*, Princeton university press, 2016.

[16] G. B. Dantzig and M. N. Thapa, *Linear programming 1: introduction*, Springer Science & Business Media, 2006.

[17] G. B. Dantzig and M. N. Thapa, *Linear programming 2: theory and extensions*, Springer Science & Business Media, 2006.

[18] J. W. Demmel, *Matrix computations (gene h. golub and charles f. van loan)*, SIAM Review, 28 (1986), pp. 252–255.

[19] L. L. Dines, *On positive solutions of a system of linear equations*, Annals of Mathematics, (1926), pp. 386–392.

[20] H. Farid, *Fundamentals of Image Processing*, `https://farid.berkeley.edu/downloads/tutorials/fip.pdf`.

[21] V. Franc, V. Hlaváč, and M. Navara, *Sequential coordinate-wise algorithm for the non-negative least squares problem*, in International Conference on Computer Analysis of Images and Patterns, Springer, 2005, pp. 407–414.

[22] X. Fu, K. Huang, N. D. Sidiropoulos, and W.-K. Ma, *Nonnegative matrix factorization for signal and data analytics: Identifiability, algorithms, and applications.*, IEEE Signal Process. Mag., 36 (2019), pp. 59–80.

[23] D. Gale, H. W. Kuhn, and A. W. Tucker, *Linear programming and the theory of games*, Activity analysis of production and allocation, 13 (1951), pp. 317–335.

[24] P. E. Gill, W. Murray, and M. H. Wright, *Practical optimization*, SIAM, 2019.

[25] P. Gingerich, *Markov analysis of cyclic alluvial sediments: Jour. sed*, 1969.

[26] H. J. Greenberg, *Klee-minty polytope shows exponential time complexity of simplex method*, University of Colorado at Denver, Denver, (1997).

[27] B. Grünbaum, V. Klee, M. A. Perles, and G. C. Shephard, *Convex polytopes*, vol. 16, Springer, 1967.

[28] E. V. Haynsworth, *On the schur complement.*, tech. report, BASEL UNIV (SWITZERLAND) MATHEMATICS INST, 1968.

[29] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge university press, 2012.

[30] M. James, *The generalised inverse*, The Mathematical Gazette, 62 (1978), pp. 109–114.

[31] N. Karmarkar, *A new polynomial-time algorithm for linear programming*, in Proceedings of the sixteenth annual ACM symposium on Theory of computing, 1984, pp. 302–311.

[32] M. Kaykobad, *Positive solutions of a class of linear systems*, Linear algebra and its applications, 72 (1985), pp. 97–105.

[33] L. G. Khachiyan, *Polynomial algorithms in linear programming*, USSR Computational Mathematics and Mathematical Physics, 20 (1980), pp. 53–72.

[34] D. Kim, S. Sra, and I. S. Dhillon, *A new projected quasi-newton approach for the nonnegative least squares problem*, Citeseer, 2006.

[35] V. Klee and G. J. Minty, *How good is the simplex algorithm*, Inequalities, 3 (1972), pp. 159–175.

[36] A. A. Kosinski, *Cramer's rule is due to cramer*, Mathematics Magazine, 74 (2001), pp. 310–312.

[37] W. C. Krumbein and M. F. Dacey, *Markov chains and embedded markov chains in geology*, Journal of the International Association for Mathematical Geology, 1 (1969), pp. 79–96.

[38] C. Lawson and R. Hanson, *Chapter 23*, Solving Least-Squares Problems. Upper Saddle River, New Jersey: Prentice-Hall, (1974).

[39] C. L. Lawson and R. J. Hanson, *Solving least squares problems*, SIAM, 1995.

[40] D. C. Lay, *Linear algebra and its applications 5th edition*, Pearson, (2016).

[41] T. L. Markham, *Nonnegative matrices whose inverses are -matrices*, Proceedings of the American Mathematical Society, 36 (1972), pp. 326–330.

[42] C. D. Meyer, *Matrix analysis and applied linear algebra*, vol. 71, Siam, 2000.

[43] K. G. Murty, *Linear programming*, Springer, 1983.

[44] Y. Nakamura, *Advanced robotics: redundancy and optimization*, Addison-Wesley Longman Publishing Co., Inc., 1990.

[45] J. M. Peña, *M-matrices whose inverses are totally positive*, Linear algebra and its applications, 221 (1995), pp. 189–193.

[46] R. J. Plemmons, *M-matrix characterizations. i—nonsingular m-matrices*, Linear Algebra and its Applications, 18 (1977), pp. 175–188.

[47] L. F. Portugal, J. J. Judice, and L. N. Vicente, *A comparison of block pivoting and interior-point algorithms for linear least squares problems with nonnegative variables*, Mathematics of Computation, 63 (1994), pp. 625–643.

[48] F. A. Potra and S. J. Wright, *Interior-point methods*, Journal of computational and applied mathematics, 124 (2000), pp. 281–302.

[49] C. R. Rao, *Generalized inverse of a matrix and its applications*, in Vol. 1 Theory of Statistics, University of California Press, 1972, pp. 601–620.

[50] S. Roman, *Positive solutions to linear systems: Convexity and separation. in: Advanced linear algebra*, (2005), p. 395–408, https://doi.org/https://doi.org/10.1007/0-387-27474-X_16.

[51] S. Roman, S. Axler, and F. Gehring, *Advanced linear algebra*, vol. 3, Springer, 2005.

[52] A. Schrijver, *Theory of linear and integer programming*, John Wiley & Sons, 1998.

[53] G. Strang, *Introduction to linear algebra. 5th edn cambridge*, United Kingdom: Wellesley-Cambridge Press.[Google Scholar], (2016).

[54] D. Tabak and B. C. Kuo, *Optimal control by mathematical programming*, SRL Publishing Company, 1971.

[55] T. Terlaky and S. Zhang, *Pivot rules for linear programming: a survey on recent theoretical developments*, Annals of Operations Research, 46 (1993), pp. 203–233.

[56] M. H. Van Benthem and M. R. Keenan, *Fast algorithm for the solution of large-scale non-negativity-constrained least squares problems*, Journal of Chemometrics: A Journal of the Chemometrics Society, 18 (2004), pp. 441–450.

[57] L. Vandenberghe and S. Boyd, *Convex optimization*, vol. 1, Cambridge University Press Cambridge, 2004.

[58] M. Wang and A. Tang, *Conditions for a unique non-negative solution to an underdetermined system*, in 2009 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton), IEEE, 2009, pp. 301–307.

[59] R. A. Willoughby, *The inverse m-matrix problem*, Linear algebra and its applications, 18 (1977), pp. 75–94.

[60] H. B. Wolfe, *Models for conditioning aging of residential structures*, Journal of the American Institute of Planners, 33 (1967), pp. 192–196.

[61] M. Wright, *The interior-point revolution in optimization: history, recent developments, and lasting consequences*, Bulletin of the American mathematical society, 42 (2005), pp. 39–56.

[62] F. Zhang, *The Schur complement and its applications*, vol. 4, Springer Science & Business Media, 2006.

[63] Y.-Y. Zhang, H.-Y. Yu, J.-K. Zhang, and J.-L. Wang, *Reliable mimo optical wireless communications through super-rectangular cover*, arXiv preprint arXiv:1607.04206, (2016).