การจัดรูปแบบคอนเวกซ์สำหรับปัญหาการวิเคราะห์เส้นทางในการจำลองสมการเชิงโครงสร้าง

นายอนุพนธ์ พฤฒิอัครวาณิช

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2559

CONVEX FORMULATIONS FOR PATH ANALYSIS PROBLEMS IN STRUCTURAL EQUATION MODELING

Mr. Anupon Pruttiakaravanich

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering Program in Electrical Engineering
Department of Electrical Engineering
Faculty of Engineering
Chulalongkorn University
Academic Year 2016

Thesis Title	CONVEX FORMULATIONS FOR PATH ANALYSIS PROBLEMS IN STRUCTURAL EQUATION MODELING
By	Mr. Anupon Pruttiakaravanich
Field of Study	Electrical Engineering
Thesis Advisor	Assistant Professor Jitkomut Songsiri, Ph.D.

Accepted by the Faculty of Engineering, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

..... Dean of the Faculty of Engineering
(Associate Professor Supot Teachavorasinskun, Ph.D.)

THESIS COMMITTEE

..... Chairman
(Professor David Banjerdpongchai, Ph.D.)

..... Thesis Advisor
(Assistant Professor Jitkomut Songsiri, Ph.D.)

..... External Examiner
(Assistant Professor Itthisek Nilkhamhang, Ph.D.)

อนุพนธ์ พฤฒิอัครวาณิช: การจัดรูปแบบคอนเวกซ์สำหรับปัญหาการวิเคราะห์เส้นทางในการจำลองสมการเชิงโครงสร้าง (CONVEX FORMULATIONS FOR PATH ANALYSIS PROBLEMS IN STRUCTURAL EQUATION MODELING)
อ. ที่ปรึกษาวิทยานิพนธ์หลัก: ผศ. ดร. จิตโกมุท ส่งศิริ, 73 หน้า

การจำลองสมการเชิงโครงสร้าง (structural equation modeling, SEM) คือหนึ่งในเทคนิคทางด้านสถิติที่ถูกนำมาใช้ในการหาโครงสร้างความสัมพันธ์เชิงสถิติของแบบจำลองหลายตัวแปร เรียกว่าการจำลองเชิงสำรวจ (exploratory modeling) หรือถูกนำมาใช้ในการทดสอบว่าแบบจำลองตามสุมมติฐานนั้นสอดคล้องกับข้อมูลที่ได้รับมาหรือไม่ เรียกว่าการจำลองเชิงยืนยัน (confirmatory modeling) การวิเคราะห์เส้นทาง (path analysis) คือปัญหาหนึ่งในการวิเคราะห์การจำลองสมการเชิงโครงสร้าง ซึ่งแบบจำลองนี้อธิบายความสัมพันธ์เชิงเหตุผลระหว่างตัวแปรที่สามารถวัดได้ในรูปแบบสมการเชิงเส้นหลายตัวแปร วิทยานิพนธ์ฉบับนี้นำเสนอการจัดรูปแบบการประมาณสองแบบสำหรับการแก้ปัญหาการวิเคราะห์เส้นทางในการจำลองสมการเชิงโครงสร้าง ในการจำลองเชิงยืนยัน รูปแบบการประมาณแบบแรกปรับเปลี่ยนเงื่อนไขจำกัดของปัญหาดั้งเดิมแบบสมการกำลังสองให้เป็นอสมการกำลังสอง การปรับดังกล่าวทำให้สามารถเปลี่ยนรูปแบบของปัญหาดั้งเดิมกลายเป็นปัญหาแบบคอนเวกซ์ซึ่งมีขั้นตอนวิธีมากมายสามารถนำมาใช้แก้ปัญหานี้ได้อย่างมีประสิทธิภาพ รูปแบบการประมาณแบบที่สองที่เพิ่มฟังก์ชันลงโทษแบบนอร์ม-1ของเมทริกซ์เส้นทางในฟังก์ชันวัตถุประสงค์ของรูปแบบการประมาณแบบแรก ส่งผลให้ได้คำตอบค่าเหมาะสุดที่มีลักษณะเบาบาง (sparse optimal solution) คำตอบค่าเหมาะสุดที่ได้จากการแก้ปัญหานี้จะมีประโยชน์ถ้าเป็นเมทริกซ์ขั้นต่ำ (low rank) ซึ่งเกิดขึ้นได้ไม่ยากในทางปฏิบัติ และเราสามารถนำคำตอบนี้ใช้เป็นตัวประมาณของเมทริกซ์ความแปรปรวนร่วมผกผันของปัญหาดั้งเดิมได้ สำหรับการแก้ปัญหาเชิงเลข วิทยานิพนธ์นี้นำเสนอการประยุกต์ใช้ขั้นตอนวิธีแบบสลับวิธีทิศทางของตัวคูณ (alternating direction method of multipliers, ADMM) ซึ่งเหมาะสมสำหรับนำมาแก้ปัญหาการประมาณสองรูปแบบในกรณีที่มีจำนวนตัวแปรมาก ผลลัพธ์ของรูปแบบการประมาณทั้งสองแบบและการแก้ปัญหาทางคณิตศาสตร์ที่มีประสิทธิภาพทำให้เราสามารถเสนอแบบแผนการหาโครงสร้างความสัมพันธ์เชิงเหตุผลระหว่างตัวแปรได้ โครงสร้างความสัมพันธ์เชิงเหตุผลที่ดีที่สุดจะถูกเลือกจากเกณฑ์การเลือกแบบจำลองห้าชนิด คือ บีไอซี (BIC), เอไอซี (AIC), เอไอซีซี (AICc), เคไอซี (KIC) และเคไอซีซี (KICc) ผลลัพธ์ที่ได้จากการจำลองแสดงให้เห็นว่า ถ้าโครงสร้างความสัมพันธ์ของแบบจำลองจริงมีลักษณะซับซ้อน เอไอซีจะให้ความแม่นยำในการเลือกแบบจำลองถูกต้องมากกว่า แต่ในทางตรงกันข้าม ถ้าโครงสร้างความสัมพันธ์ของแบบจำลองจริงมีลักษณะเบาบาง บีไอซี, เอไอซีซีและเคไอซีซีจะให้ความแม่นยำในการเลือกแบบจำลองได้ถูกต้องมากกว่าเกณฑ์อื่น ในส่วนของการประยุกต์ใช้งานกับข้อมูลจริง วิทยานิพนธ์นี้ทดลองหาโครงสร้างความสัมพันธ์เชิงเหตุผลระหว่างพื้นที่ในสมองจากข้อมูลภาพเรโซแนนซ์แม่เหล็กเชิงหน้าที่หรือเอ็ฟเอ็มอาร์ไอ (functional magnetic resonance imaging, fMRI) ที่ถูกบันทึกจากการทดลองแบบการกระตุ้นด้วยการมองเห็นแล้วจึงเคลื่อนไหวมือ ผลลัพธ์ที่ได้แสดงให้เห็นว่า พื้นที่โซมาโตเซนซอรี (somatosensory), พาเรียทัล (parietal), พรีมอเตอร์ (premotor) และมอเตอร์ (motor) มีความสัมพันธ์อย่างเห็นได้ชัด โดยเฉพาะอย่างยิ่งคู่โซมาโตเซนซอรี→วิชวล (visual), โซมาโตเซนซอรี→พาเรียทัล และโซมาโตเซนซอรี→พรีมอเตอร์ มีความสัมพันธ์ชัดเจนมากที่สุดสามคู่

ภาควิชา วิศวกรรมไฟฟ้า ..... **ลายมือชื่อนิสิต** .....
สาขาวิชา วิศวกรรมไฟฟ้า ..... **ลายมือชื่อ อ.ที่ปรึกษา** .....
ปีการศึกษา 2559 .....

5770348021 : MAJOR ELECTRICAL ENGINEERING

KEYWORDS : STRUCTURAL EQUATION MODELING / PATH ANALYSIS / CONFIRMATORY SEM / EXPLORATORY SEM / CONVEX PROGRAM

ANUPON PRUTTIAKARAVANICH : CONVEX FORMULATIONS FOR PATH ANALYSIS PROBLEMS IN STRUCTURAL EQUATION MODELING.
ADVISOR : ASST. PROF. JITKOMUT SONGSIRI, Ph.D., 73 pp.

Structural equation modeling (SEM) is a statistical technique used for seeking a statistical causal multivariate model (called exploratory modeling) or for testing whether the model is supported by the given data (called confirmatory modeling). Path analysis is a problem in SEM analysis where its model describes causal relations among measured variables in a form of multivariable linear equations. This thesis proposes two alternative estimation formulations for solving problems of path analysis in SEM. For confirmatory SEM, our first formulation relaxes the original nonlinear equality constraints of the model parameters to an inequality, allowing us to transform the original problem into a convex problem that can be solved by many existing efficient algorithms. The second formulation is a regularized estimation proposed for exploratory SEM by adding ℓ_1-type penalty of the path matrix into the cost objective of the first formulation which leads to sparse solutions. Practically, our optimal solution is useful when it has low rank which occurs under a mild condition on problem parameters. This solution can be used as an estimate of the inverse of covariance matrix from the original problem. Another contribution of this thesis is a numerical method based on ADMM algorithm that is suitable for solving the two formulations in a large-scale setting. This thesis also provides a scheme of learning a causal structure among variables by applying both proposed formulations. The best causal structure from our scheme is chosen from five model selection criterions, those are BIC, AIC, AICc, KIC and KICc. Our approach is examined with simulated and real data sets. The simulation results show that if the causal structure of true model is complex, AIC provides the better accuracy while BIC, AICc and KICc yield better performance when the causal structure of true model is simpler. An application of this scheme has been preliminarily illustrated by learning causal relations among brain regions from fMRI data, recorded from visual-hand hemifield stimuli experiments. A brain network from our findings shows strong relations among somatosensory, parietal, premotor, and motor area. In particular, the dominant pairs of strong connection are somatosensory→visual, somatosensory→parietal and somatosensory→premotor.

Department : Electrical Engineering
Field of Study : Electrical Engineering
Academic Year : 2016

Student's Signature .....
Advisor's Signature .....

Acknowledgements

I would like to express sincere thanks to my thesis advisor, Assistant Professor Jitkomut Songsiri for her invaluable help and constant encouragement throughout the course of this research. I am most grateful for her teaching and advice, not only the research methodologies but also many other methodologies in life. I would not have achieved this far and this thesis would not have been completed without all the continuous support that I have always received from her.

I am grateful to all committees, Professor David Banjerdpongchai serving as a chairman and Assistant Professor Itthisek Nilkhamhang serving as external examiner, who devote valuable times for my proposal exam and thesis defense.

In addition, I would like to thank Department of Electrical Engineering of Chulalongkorn University for the financial support throughout my master's study by providing me the teacher assistant (TA) and research assistant (RA) scholarships.

Finally, I am grateful for the support from my parents and my friends throughout the period of this research.

Contents

Page

CHAPTER Page

List of Figures

CHAPTER I

INTRODUCTION

1.1 Introduction

Structural equation modeling (SEM) is a statistical technique used for seeking a statistical causal multivariate model (called exploratory modeling) or for testing whether the model are supported by the given data (called confirmatory modeling). In the model, it includes two types of variables that are assumed to be random, *observed* variables which can be directly measured and the *latent* variables that cannot be exactly measured. The relationship between these two types is explained by a linear model where a nonzero coefficient of a latent variable explain a cause or influence from such variable to an observed variable. SEM has a long history since 1980s and is widely used in many behavioral researches such as in psychology [1], sociology [2, 3], business [4], and many more; a history background can be found in [5, §1]. Path analysis is a special problem in SEM where it provides a model for explaining relationships among measured variables only (no latent variables included in the model) and additive error terms. In scientific research, observed variables are often of primary interest. For example, one aims to explore causal relationship among brain regions from brain signals (such as fMRI data) [6–10] where the entry of the path coefficient matrix in the model explains how much change in the activity of one region influences another region.

The problem in path analysis starts from constructing a *hypothetical model* where directional paths from one variable to another variable are assumed from a prior knowledge about relationship structure of variables of interest. This can be encoded as the zero structure of the *path coefficient* matrix in the model and becomes a part of problem constraint. The formulation is then to estimate the value of nonzero entries in the path matrix and the covariance matrix of model residual errors so that the model-reproduced covariance matrix fits well with the sample covariance matrix in an optimal sense, evaluated by various types of criterion functions such as maximum likelihood, ordinary or weighted least-squares [5, §4]. When the zero structure of the path matrix is hypothetically given, the resulting estimation problem is called *confirmatory SEM* which find many applications in behavioral research. In contrast, one may seek for a zero structure of the path matrix that best fits the data since its pattern reveals a causal structure of variables such as a problem of learning brain connectivity in neuroscience. The latter type of estimation problem is referred to as *exploratory SEM*. An existing approach for latter problem is to begin with a base model where a certain set of paths are affirmative but the existence of some other paths is in question. This results in a set of a few candidate models associated with different zero structures of the path matrix and the significance of the difference between these models can be determined from the χ^2 statistic [5, §7]. Examples of this approach

can be seen in brain network study [6, 7] where only a few variables (in the order up to 10 brain regions) are selected. One can locally search for a path structure by starting from a null model (all path coefficients are zero) and sequentially allowing the coefficient corresponding to the largest Lagrangian multiplier to be nonzero [8]. The most optimal but far from feasible approach is to perform an exhaustive search that enumerates all possible path pattern with a fixed number of paths and chooses the model corresponding to the lowest minimized maximum likelihood function [10]. It is known that the number of all possible models grows exponentially to the number of variables, so it is not feasible as the number of variables increases.

Both confirmatory and exploratory SEM problems are nonlinear optimization problems in matrix variables with quadratic equality and positive definite cone constraints. Many existing SEM commercial softwares have been developed such as LISREL, EQS, Mplus [5, 11, 12]. These softwares implement iterative methods such as Newton-Raphson, or gradient descent to estimate the model parameters [13, §7], [5, §4], so a starting value for the update iteration is required. Though these numerical methods work well under normal conditions, it is also known that some initial values may not lead to the convergence in the optimal solution or may stuck into a local minima, hence several strategies for selecting initial values have been proposed [5, §4]. These include choosing an instrumental variable estimate or selecting the strength of the path coefficient magnitude. When the iterative method in these softwares does not converge, the user is suggested not to interpret the result.

In this work, we present two alternative estimation formulations for both confirmatory and exploratory SEM problems. In addition, the original nonlinear equality constraints of the model parameters is relaxed to an inequality, allowing us to transform the problem into convex formulations that are supported by rich duality theory for analyzing the problem properties and lead to many existing convex program solvers. For exploratory SEM, we propose a formulation whose objective function is added with an ℓ_1-type regularization of the path coefficient matrix, called sparse SEM. This is a known result that such formulation is regarded as a *lasso* formulation [14] and doing so encourages many zeros in the path matrix solution, allowing us to read off the zero pattern and interpret it as a causal structure of the variables. We solve our estimation formulations by a first-order method, called alternating direction method of multipliers or ADMM. This method requires a feasible amount of memory storage suitable for large-scale implementation and the main computational cost in each update step depends on eigenvalue decomposition of a symmetric matrix with size $2n$ where n is the number of variables. We also show that, under a condition on problem parameters of both confirmatory and sparse SEM, our optimal covariance error is diagonal, meaning that errors are uncorrelated, and the optimal solution has low rank, providing an estimate of the path matrix for the original problem.To apply our estimation formulation to real-world applications, we explore causal relations among brain regions from fMRI data.

Despite a difference in our estimation formulation and the original one, we believe that our proposed formulations serve two folds. Firstly, unlike previous SEM applications that only a few variables are of interest, many applications tend to consider a much larger number of variables such as fMRI studies where the variables are neuronal activities and its number is up to thousand [15]. Existing approaches of learning causal structures in the exploratory SEM may experience a computational difficulty in terms of memory storage or convergence. Secondly, our solution for confirmatory SEM is obtained under an assumption of homoskedasticity of residual errors, so if this assumption holds, ours and the original solution coincide. Even if it does not hold, so our solution is not optimal for the original problem but it can be served as a starting value for the iterative algorithm used in the original one in case that the convergence is not obtained.

1.2 Objectives

1. To explore causal relationships among observed variables by using confirmatory and exploratory Structural Equation Modeling (SEM).

2. To present alternative estimation formulations for special case problems in path analysis in a convex framework.

3. To provide efficient algorithms for solving our alternative estimation problems in large-scale settings, in order to apply to real-world applications.

1.3 Scope of Thesis

1. Propose two alternative convex formulations for the problems of confirmatory SEM and exploratory SEM. Express the dual problems and KKT conditions of such formulations.

2. Provide the conditions on problem parameters that the optimal solution in our estimation formulation is useful and can be used as the inverse of estimated covariance matrix of the original problem. This solution is referred to as a *low rank* solution.

3. Based on the use of two formulations, we provide a scheme for learning causal relation structures in variables.

4. Provide an efficient implementation of gradient-based methods for solving our estimation problems in a larger scale that what have been done in the literatures.

5. Apply our alternative estimation formulations to real applications. For instance, we explore causal relations among brain regions via functional magnetic resonance imaging (fMRI) data.

1.4 Methodology

1. A convex formulation for the confirmatory SEM is obtained by relaxing the original equality constraint into an inequality. The resulting problem becomes a semidefinite programing and it is referred to as *primal convex SEM*.

2. Useful solutions of the exploratory SEM are sparse. Therefore, we apply ℓ_1-norm minimization and introduce ℓ_1-penalty to the primal convex SEM.

3. The relations between the original problem and our formulation can be explained under some conditions on the problem parameters. We apply Farka's lemma and use KKT conditions for this purpose.

4. Alternating direction method of multipliers (ADMM) or other optimal fast gradient methods are implemented to solve the problems in large scale.

1.5 Expected Outcomes

1. Two estimation formulations cast in a convex framework.

2. Efficient algorithms for solving our estimation formulations in large-scale settings.

1.6 Achievements

The contributions of this thesis are as follows:

- We propose two alternative convex formulations for both confirmatory and exploratory SEM. For confirmatory SEM, we relax an equality constraint of the original problem and we refer this proposed formulation as primal convex SEM. To learn a sparse causal relation structure among the variables, we add the ℓ_1-regularization term on the objective function of primal convex SEM, which we refer this proposed formulation as sparse SEM.

- We provide an implementation of alternating direction method of multipliers (ADMM) for solving our two estimation formulations in a large-scale framework.

- We provide a scheme for learning causal relations among observed variables which is a combination between two proposed formulations, primal convex SEM and sparse SEM. This scheme includes the model selection procedure for choosing the best model.

1.7 Thesis Outline

Our thesis is organized as follows. Chapter 2 summarizes the mathematical formulation of the original path analysis problem which is the maximum likelihood estimation with an quadratic equality constraint. Chapter 3 describes our convex formulation for confirmatory SEM, referred to as primal convex SEM, and shows that the solution can be further used under the condition of having a low rank solution at optimum. Another convex formulation for exploring a sparse causal relation among variables is proposed in chapter 4 where an ℓ_1-regularization is introduced in the cost objective of primal convex SEM, referred to as the sparse SEM. We show that sparse solutions are obtained and the sparsity can be controlled by a regularization parameter. Applying the two formulations is used in the exploratory SEM, proposed in 5. Numerical methods for solving our proposed formulations are explained in chapter 6. This is the alternating direction method of multipliers method or ADMM which is suitable for solving large-scale convex problems and yield a reasonably good rate of convergence. Numerical experiments in chapter 7 describes simulated examples under the condition that low rank solutions are obtained, the effect of our problem parameters on such conditions and some important results of our sparse and exploratory SEM. The performance of numerical methods for solving our estimation formulations are also included to this chapter. Lastly, chapter 8 concludes our work and comments. Our proofs, the derivation of dual problems of our formulations, the derivation of the critical value of regularization parameters, and the miscellaneous proofs are shown in Appendices, which could be omitted if the reader is familiar with the duality theory.

Notation. $\mathbf{S}^n$ denotes the set of symmetric matrices of size $n \times n$ and $\mathbf{S}^n_+$ denotes the set of positive semidefinite matrices of size $n \times n$. For a square matrix X, $\mathbf{tr}(X)$ is the trace of X and $\mathbf{diag}(X)$ is a diagonal matrix containing diagonal entries of X.

CHAPTER II

BACKGROUND ON PATH ANALYSIS IN SEM

Structural equation modeling (SEM) starts with a set of variables involved in a study, measured variables and latent variables. Measured variables are simply the ones that can be directly measured (physical quantities), while latent variables are variables that cannot be directly (or exactly) measured such as intelligence, attitude, etc. Each of these variables can be regarded as either endogenous or exogenous. An endogenous variable gets an influence from others while an exogenous variable affects the other variables. A general mathematical model in SEM explains a linear relationship from latent variables to measured variables and also includes error terms of each variable [5, 13, 16].

In contrast to research in social science, we are only interested in application of SEM that involves only with observable variables. For this reason, we focus on a special class of model in SEM that is described by a multiple linear regression:

$$Y = c + AY + \epsilon \tag{2.1}$$

where $Y \in \mathbf{R}^n$ is the measured (or observed) variables, $c \in \mathbf{R}^n$ is a constant vector representing a baseline, and $\epsilon \in \mathbf{R}^n$ is the model error, assumed to be Gaussian distributed. The matrix $A \in \mathbf{R}^{n\times n}$ denotes the path matrix where A_{ij} represents a causal relationship among variables in the model, *i.e.*, if $A_{ij} = 0$ then there is no path from Y_j to Y_i. In other words, a pattern of nonzero entries in A reveals a causal structure of variables in the model. If this structure is assumed from a *prior* knowledge, then the problem of estimating A is called *confirmatory SEM*.

Let S be a sample covariance matrix of Y which can be computed from sample measurements of Y. Let Σ be the model-reproduced covariance matrix of Y, derived from (2.1)

$$\Sigma = (I - A)^{-1}\Psi(I - A)^{-T} \tag{2.2}$$

where $\Psi = \mathbf{cov}(\epsilon)$. The estimation in SEM is to seek for A and Ψ such that the estimated Σ is close to S in the sense that the Kullback-Leiber divergence function

$$d(S, \Sigma) = \log\det\Sigma + \mathbf{tr}(S\Sigma^{-1}) - \log\det S - n$$

is minimized, while Σ, A and Ψ are related by (2.2). Moreover, the structure of the path matrix is presumably encoded by a model hypothesis: i) $A_{ij} = 0$ if there is no link from Y_j to Y_i and ii) we always have $\mathbf{diag}(A) = 0$, meaning that there is no path from Y_i to itself. To specify the zero structure of A, we then define the associated index set $I_A \subseteq \{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\}$ with properties that i) $(i, j) \in I_A$ if $A_{ij} = 0$ and ii) $\{(1, 1), (2, 2), \ldots, (n, n)\} \subseteq I_A$ since $\mathbf{diag}(A) = 0$. In short, I_A denotes the index set of *hypothetical zero entries* in A and it must include the diagonal entries.

Given the index set I_A, we define a *projection operator* $P : \mathbf{R}^{n\times n} \to \mathbf{R}^{n\times n}$

$$P(X) = \begin{cases} X_{ij}, & (i,j) \in I_A, \\ 0, & \text{otherwise}, \end{cases} \tag{2.3}$$

and denote

$$P^c = I - P. \tag{2.4}$$

The operators P^c and P are both self-adjoint, *i.e.*, $\mathbf{tr}(Y^T P(X)) = \mathbf{tr}(P(Y)^T X)$ and that $P^c(P(X)) = 0$. Define $Q : \mathbf{S}^{2n} \to \mathbf{S}^{2n}$ a linear operator given by

$$Q\left(\begin{bmatrix} X_1 & X_2^T \\ X_2 & X_4 \end{bmatrix}\right) = \begin{bmatrix} 0 & P^c(X_2^T) \\ P^c(X_2) & 0 \end{bmatrix}, \tag{2.5}$$

where $X_1, X_4 \in \mathbf{S}^n$ and $X_2 \in \mathbf{R}^{n\times n}$. From the definitions of P and Q, we note that $P(X)$ just extracts the entries X_{ij} for (i,j) belonging to I_A, while $Q(X)$ are all zero except that the $(1,2)$ block of $Q(X)$ contains nonzero (i,j) entries for $(i,j) \in I_A$. These two projection operators will be used repeatedly in our analysis.

With the definition of P and a change of variable $X = \Sigma^{-1}$, the estimation problem corresponding to confirmatory SEM is

$$\begin{array}{ll} \text{minimize} & -\log\det X + \mathbf{tr}(SX) - \log\det S - n, \\ \text{subject to} & X = (I-A)^T \Psi^{-1}(I-A), \\ & P(A) = 0, \end{array} \tag{2.6}$$

with variables $A \in \mathbf{R}^{n\times n}$, $\Psi \in \mathbf{S}_+^n$ and $X \in \mathbf{S}_+^n$. The condition $P(A) = 0$ basically explains the zero constraint on the entries of A, and when there is no information on the path matrix, this condition reduces to $\mathbf{diag}(A) = 0$. The problem (2.6) is one of estimation formulations considered in SEM context [5, §4]. Other cost objectives are also used such as ordinary or weighted least-squares.

Special case. If the constraint $P(A) = 0$ reduces to $\mathbf{diag}(A) = 0$ (we allow A to have as many free parameters as possible), then we can make the cost objective to be zero by solving $S^{-1} = (I-A)^T\Psi^{-1}(I-A)$ where S is given while A and Ψ are free variables. In this case, we can arbitrarily make Ψ diagonal. In other words, one can always find a factor B with $\mathbf{diag}(B) = 1$ and a diagonal D such that $S^{-1} = B^T D B$. Such factors can be obtained by simply performing an eigenvalue decomposition of S^{-1} and normalize the matrix of eigenvectors to have unit diagonals. Another way is to perform LDL^T decomposition where L can be normalized to have unit diagonals. In this case, the optimal path matrix to (2.6) is not unique; one can obtain A as dense or lower triangular matrix. This could be problematic if one would like to read a causality structure from the zero pattern in the estimated A. For this reason, it is common to assume some structure in A and diagonal structure in Ψ (meaning the error terms are uncorrelated). Specifically, we define the degree

of freedom (df) by,

$$\mathrm{df} = \text{ the number of known parameters } - \text{ the number of estimated parameters } . \tag{2.7}$$

Referring to the cost function and constraints in (2.6), the number of known parameters is the number of entries in the sample covariance matrix and is equal to $n(n-1)/2$ where n is the number of observed variables. The number of free parameters in (2.7) is the total number of entries in A plus the total number of entries in Ψ. One can use df as a guideline for identifying the uniqueness of solution. When df is negative, the estimator may not be unique. We say that the model is *identifiable* if the df is nonnegative [11, p. 35]. Therefore, to find a unique solution in a path analysis problem, we must have some assumptions on the path matrix A and noise covariance Ψ to attain the nonnegative values in df as a necessary condition. For example, if we assume that Ψ is diagonal, then we must assume the percentage of zero entries in A about 50% to attain zero df.

CHAPTER III

CONVEX FORMULATION FOR CONFIRMATORY SEM

The problem (2.6) is obviously nonconvex due to the quadratic equality constraint. In this chapter, we propose an alternative convex formulation and its dual problem. We consider a special case of path analysis problem where the covariance error is allowed to be diagonal, meaning that the residual errors are assumed to be uncorrelated. The solution to our formulation is useful only when it is low rank at optimum which will be shown to occur under some condition on a problem parameter. The solutions to our formulation and the original problem agree when the covariance error is specified to be a multiple of the identity matrix.

Consider a convex relaxation of the constraint (2.2) to $X \succeq (I-A)^T\Psi^{-1}(I-A)$ which is equivalent to

$$\begin{bmatrix} X & (I-A)^T \\ I-A & \Psi \end{bmatrix} \succeq 0$$

by using the Schur complement. We then propose an alternative convex formulation:

$$\begin{array}{ll} \text{minimize} & -\log\det X + \mathbf{tr}(SX), \\ \text{subject to} & \begin{bmatrix} X & (I-A)^T \\ I-A & \Psi \end{bmatrix} \succeq 0, \\ & 0 \preceq \Psi \preceq \alpha I, \\ & P(A) = 0, \end{array} \tag{3.1}$$

with variables $X \in \mathbf{S}^n, A \in \mathbf{R}^{n\times n}$ and $\Psi \in \mathbf{S}^n$, where $\alpha > 0$ is a given parameter. We note that the inequality constraint $\Psi \preceq \alpha I$ is additionally introduced to prevent (3.1) from having a trivial solution, *e.g.*, Ψ can be arbitrarily large and $A = 0$. We justify that α can serve as an upper bound on the covariance error in SEM. Throughout this thesis, we refer to (3.1) as *the primal convex SEM formulation* which falls into a type of semidefinite programming. We can see that for a given α, a numerical solution can be solved by many existing convex program solvers. One example is `CVX` which is a MATLAB package for specifying and solving convex programs [17].

If we define a variable

$$X = \begin{bmatrix} X_1 & X_2^T \\ X_2 & X_4 \end{bmatrix}, \quad X_4 = \Psi, \quad X_2 = I - A,$$

we see that $P(X_2) = P(I) - P(A) = P(I) - 0 = I$ (note that the P projects the entries assigned by I_A which includes the diagonal terms). Another equivalent formulation of the primal is

$$\begin{array}{ll} \text{minimize} & -\log\det X_1 + \mathbf{tr}(SX_1), \\ \text{subject to} & X = \begin{bmatrix} X_1 & X_2^T \\ X_2 & X_4 \end{bmatrix} \succeq 0, \\ & 0 \preceq X_4 \preceq \alpha I, \\ & P(X_2) = I, \end{array} \tag{3.2}$$

with variable $X \in \mathbf{S}^{2n}$ where $X_1, X_4 \in \mathbf{S}^n$ and $X_2 \in \mathbf{R}^{n\times n}$.

The dual problem of (3.1) is

$$\begin{array}{ll} \text{minimize} & -\log\det(S - Z_1) - 2\,\mathbf{tr}(Z_2) - \alpha\,\mathbf{tr}(Z_4) + n, \\ \text{subject to} & Z = \begin{bmatrix} Z_1 & Z_2^T \\ Z_2 & Z_4 \end{bmatrix} \succeq 0, \\ & Q(Z) = 0, \end{array} \tag{3.3}$$

with variable $Z \in \mathbf{S}^{2n}$ where each block Z_k has size $n \times n$. The constraint $Q(Z) = 0$ explains that i) Z_1 and Z_4 are freely nonzero, ii) the corresponding entries of block Z_2 to the zero entries in A are free, otherwise they are all zeros. If the condition $P(A) = 0$ reduces to $\mathbf{diag}(A) = 0$ in the primal problem, then $Q(Z) = 0$ in the dual is simplified to that Z_2 is diagonal. Details of dual problem are shown in Appendix 9.1.

Problem assumptions. From the cost function in (3.1), we will show that S must be positive definite. Otherwise, the problem could be unbounded below. To show this, assume S has the eigenvalue decomposition $S = UDU^T$. Then it follows that $\mathbf{tr}(SX) = \mathbf{tr}(UDU^TX) = \mathbf{tr}(DU^TXU)$. Let $Y = U^TXU$ and since $\det X = \det Y$, we can write $f(X) = f(Y) = -\log\det Y + \mathbf{tr}(DY)$. If S is positive semidefinite, then $d_{ii} = 0$ for some i, and we can choose Y to be diagonal where y_{ii} is chosen to be arbitrarily large. Hence, the term $\mathbf{tr}(DY) = 0$ but $-\log\det Y \to -\infty$, leading the cost function to be unbounded below.

3.1 KKT conditions

The KKT conditions are derived as the optimality condition for the optimal solution to (3.1). These conditions are:

Zero gradient of the Lagrangian

$$X = (S - Z_1)^{-1}. \tag{3.4}$$

Primal feasibility

$$\begin{aligned} (I - A)^T \Psi^{-1} (I - A) &\preceq X, && (3.5)\\ 0 \prec \Psi &\preceq \alpha I, && (3.6)\\ P(A) &= 0. && (3.7) \end{aligned}$$

Dual feasibility

$$\begin{aligned} Z &\succeq 0, && (3.8)\\ Q(Z) &= 0. && (3.9) \end{aligned}$$

Complementary Slackness condition

$$\begin{aligned} \mathbf{tr}\left(\begin{bmatrix} Z_1 & Z_2^T \\ Z_2 & Z_4 \end{bmatrix} \begin{bmatrix} X & (I-A)^T \\ I - A & \Psi \end{bmatrix} \right) &= 0, && (3.10)\\ \mathbf{tr}(Z_4(\Psi - \alpha I)) &= 0. && (3.11) \end{aligned}$$

We will use these conditions to analyze the solution properties later throughout this thesis.

3.2 Trivial solutions

In this section, we show that there is a critical value α_c such that if the optimal dual $Z^\star = 0$ then $\alpha \geq \alpha_c$, *i.e.*, if the trivial solution in the dual occurs then we have used too large value of α. From the zero gradient of the Lagrangian condition and the primal feasibility, if $Z = 0$ then

$$X = S^{-1}, \quad X \succeq (I - A)^T \Psi^{-1} (I - A),$$

which means Ψ can be sufficiently large and the RHS of the above inequality can be sufficiently small. The matrix $X = S^{-1}$ can be chosen to be greater than $(I - A)^T \Psi^{-1} (I - A)$ as desired.

This section presents an important result that the critical value α_c turns out to be the *harmonic mean of the eigenvalues of* S. To show this, we firstly make a change of variable: $\tilde{A} = A - P(A)$ so that $P(\tilde{A}) = 0$. If $Z^\star = 0$, then the KKT conditions are reduced to

$$\begin{aligned} X = S^{-1}, \quad 0 \preceq \Psi \preceq \alpha I,\\ S^{-1} \succeq (I - \tilde{A})^T \Psi^{-1} (I - \tilde{A}), \end{aligned} \tag{3.12}$$

which is a feasibility problem. We can equivalently show that if $\alpha \leq \alpha_c$ then (3.12) has no solution. To this end, we will prove by contradiction: if $\alpha \leq \alpha_c$ and then (3.12) has a solution by applying a generalization of Farka's lemma to semidefinite programming [18].

Lemma 1. *The system*

$$Z \succeq 0, \quad \mathbf{tr}(GZ) > 0, \quad \mathbf{tr}(F_i Z) = 0, i = 1, 2, \ldots, n$$

is a strong alternative for the nonstrict LMI:

$$\sum_{i=1}^{n} x_i F_i + G \preceq 0,$$

if the matrices F_i satisfy $\sum_{i=1}^{n} v_i F_i \succeq 0$ implies that $\sum_{i=1}^{n} v_i F_i = 0$.

Our result is stated in the following proposition.

Proposition 2. *Let $\alpha_c = n/\,\mathbf{tr}(S^{-1})$ (the harmonic mean of the eigenvalues of $S \succ 0$). If $\alpha \leq \alpha_c$ then* (3.12) *has no solution, i.e., $Z = 0$ cannot be an optimal solution for the dual problem* (3.3)*.*

Proof. The feasibility problem (3.12) can be expressed as an LMI

$$\begin{bmatrix} S^{-1} & (I - \tilde{A})^T & 0 \\ I - \tilde{A} & \Psi & 0 \\ 0 & 0 & \alpha I - \Psi \end{bmatrix} \succeq 0 \tag{3.13}$$

or equivalently, $G + \sum_{ij} A_{ij} F_{ij} + \sum_{ij} \Psi_{ij} H_{ij} \preceq 0$ where

$$G = \begin{bmatrix} -S^{-1} & -I & 0 \\ -I & 0 & 0 \\ 0 & 0 & -\alpha I \end{bmatrix}, \quad \sum_{ij} A_{ij} F_{ij} = \begin{bmatrix} 0 & \tilde{A}^T & 0 \\ \tilde{A} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\sum_{ij} \Psi_{ij} H_{ij} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\Psi & 0 \\ 0 & 0 & \Psi \end{bmatrix}.$$

We note that A_{ij} and Ψ_{ij} are the (i, j) entries of $\tilde{A}$ and Ψ, respectively. The matrices F_{ij} and H_{ij} are the common choice of standard basis matrices that make up to the above summations. To describe more details, let E_{ij} be a standard basis matrix for set of $n \times n$ matrices with zero diagonals and S_{ij} be a standard basis matrix for $\mathbf{S}^n$. In other words, the entries of E_{ij} are all zero except that the (i, j) entry is 1. Similarly, the entries of S_{ij} are all zero except that the (i, j) and (j, i) entries are 1. The expressions of F_{ij} and H_{ij} are

$$F_{ij} = \begin{bmatrix} 0 & E_{ij}^T & 0 \\ E_{ij} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \text{ for } (i, j) \notin I_A,$$

$$H_{ij} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -S_{ij} & 0 \\ 0 & 0 & S_{ij} \end{bmatrix}, \text{ for } i \geq j = 1, 2, \ldots, n.$$

From Lemma 1, the LMI (3.13) has no solution if and only if $\exists U \succeq 0, U \neq 0$ such that

$$\mathbf{tr}(GU) \geq 0, \quad \mathbf{tr}(F_{ij}U) = 0, \ \text{ for } (i,j) \notin I_A,$$
$$\mathbf{tr}(H_{ij}U) = 0, \ \text{ for } i \geq j.$$

In what follows, we will show that there always exists such matrix U under the condition $\alpha \leq \alpha_c$. For scalars γ and β with $\beta \geq 0$ and $\gamma \neq 0$, we construct a positive semidefinite matrix U of the form

$$U = \begin{bmatrix} (\gamma^2/\beta)I & \gamma I & 0 \\ \gamma I & \beta I & 0 \\ 0 & 0 & \beta I \end{bmatrix}.$$

With this choice, we can easily check that $\mathbf{tr}(F_{ij}U) = 0$ regardless of I_A (as long as I_A contains the indices of diagonal entries of A), and that $\mathbf{tr}(H_{ij}U) = 0$. We also see that

$$\sum_{ij} A_{ij}F_{ij} + \sum_{ij} \Psi_{ij}H_{ij} = \begin{bmatrix} 0 & \tilde{A}^T & 0 \\ \tilde{A} & -\Psi & 0 \\ 0 & 0 & \Psi \end{bmatrix} \succeq 0$$

implies that $\Psi = 0$ and consequently conclude that $\tilde{A} = 0$ because 0 is in the leading $(1,1)$ block. Lastly, the condition $\mathbf{tr}(GU) \geq 0$ is expressed as

$$\frac{\mathbf{tr}(S^{-1})}{\beta}\left(\gamma^2 + \frac{2n\beta}{\mathbf{tr}(S^{-1})}\gamma + \frac{n}{\mathbf{tr}(S^{-1})}\alpha\beta^2\right) \leq 0. \tag{3.14}$$

The above quadratic polynomial in γ can be expressed in terms of α and α_c as

$$\gamma^2 + 2\alpha_c\beta\gamma + \alpha\alpha_c\beta^2 \leq 0.$$

Therefore, if $\alpha \leq \alpha_c$ then we can always choose any negative real value of γ in the interval

$$\left(-\alpha_c\beta(1+\sqrt{1-\alpha/\alpha_c}), -\alpha_c\beta(1-\sqrt{1-\alpha/\alpha_c})\right)$$

so that (3.14) is satisfied. This concludes that if $\alpha \leq \alpha_c$ the alternative of (3.13) always has a solution. This completes the proof.

3.3 Low rank solutions of the primal convex SEM

The solution of the primal convex SEM is useful if $X = (I-A)^T\Psi^{-1}(I-A)$ at optimum as we can use X as the estimate of Σ^{-1}. This occurs if and only if the rank of

$$\begin{bmatrix} X & (I-A)^T \\ (I-A) & \Psi \end{bmatrix}$$

is n at optimum. Therefore, we aim to find a relation between the parameter α and the low rank optimal solution of (3.1) from the complementary slackness condition. The result in section 3.2 gave us a hint that if α is too large, then $\mathbf{rank}(X) > n$ which is to be avoided.

To show this in detail, we refer to the complementary slackness condition (3.10) and from a property of trace: $\mathbf{tr}(AB) = 0 \iff AB = 0$ for $A, B \succeq 0$, we have

$$\underbrace{\begin{bmatrix} Z_1 & Z_2^T \\ Z_2 & Z_4 \end{bmatrix}}_{Z} \underbrace{\begin{bmatrix} X & (I-A)^T \\ I-A & \Psi \end{bmatrix}}_{W} = 0. \tag{3.15}$$

The result in (3.15) further shows that the columns of W are in the nullspace of Z. Therefore, we have $\mathbf{rank}(W) = \mathbf{nullity}(Z)$ and that $\mathbf{rank}(Z) = 2n - \mathbf{rank}(W)$. Since $X \succ 0$ is an implicit constraint, this implies that X must be full rank, *i.e.*, the $(1,1)$ block of W has rank n. The rank of W must satisfy $n \leq \mathbf{rank}(W) \leq 2n$ and therefore $0 \leq \mathbf{rank}(Z) \leq n$.

We obtain a *low rank* solution when the optimal primal of (3.1) and the optimal dual of (3.3) satisfies

$$X = (I-A)^T \Psi^{-1} (I-A) \text{ or equivalently } \mathbf{rank}(Z) = n,$$

(because $\mathbf{rank}(W) = n$). Furthermore, when this holds, $\mathbf{rank}(Z_4) = n$ and from the complementary slackness condition (3.11), it gives $\Psi = \alpha I$, *i.e.*, the estimated covariance error becomes a diagonal matrix. From section 3.2, we have shown that if α is smaller than $\alpha_c = n/\mathbf{tr}(S^{-1})$, then the optimal dual solution is not zero. This suggests that we can consider three ranges of α where the rank of Z varies as shown in Figure 3.1. The value of α_c lies some where in the interval that results in $0 < \mathbf{rank}(Z) < n$.

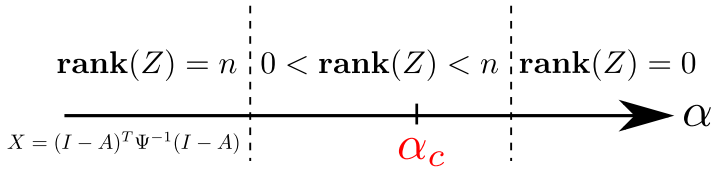


Figure 3.1: Possibility of $\mathbf{rank}(Z)$ as α varied.

We can also show that the minimum eigenvalue of S lies on the left to the harmonic mean of the eigenvalues of S. Suppose $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ are eigenvalues of S. It follows that

$$\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \cdots + \frac{1}{\lambda_n} \leq \frac{n}{\lambda_1}.$$

Since the trace of a matrix is the sum of its eigenvalues, we have $\mathbf{tr}(S^{-1}) = \sum_{k=1}^{n} 1/\lambda_k$ and this further implies that

$$\alpha_c = \frac{n}{\mathbf{tr}(S^{-1})} = \frac{n}{\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \cdots + \frac{1}{\lambda_n}} \geq \lambda_1 = \lambda_{\min}(S).$$

If $\alpha = \lambda_{\min}(S)$, then it is often the case that $\mathbf{rank}(Z) = n$ which will be shown in the numerical experiment section. This suggests us that we have put a constraint $\Psi \preceq \lambda_{\min}(S) I \preceq S$ into the estimation problem. Our justification is that we control the covariance error to be less than the covariance of the variables.

3.4 Uniqueness of solution

In this section, we attempt to derive a sufficient condition of the uniqueness of solution for the primal convex SEM. However, we have not successfully found a practical condition and leave this as an open problem. The information in this section is therefore presented as a background supplement for future study. Firstly, we assume that α is small enough so that the two matrices Z and W in the complementary slackness condition:

$$\underbrace{\begin{bmatrix} Z_1 & Z_2^T \\ Z_2 & Z_4 \end{bmatrix}}_{Z} \underbrace{\begin{bmatrix} X & (I-A)^T \\ I-A & \Psi \end{bmatrix}}_{W} = 0, \tag{3.16}$$

have low rank. When low rank solutions in Z and W are obtained, the KKT conditions of the primal convex SEM are

$$\begin{aligned} X &= (S - Z_1)^{-1} = \frac{1}{\alpha}(I-A)^T(I-A), && (3.17)\\ P(A) &= 0, && (3.18)\\ Z_4 &= Z_2 Z_1^{-1} Z_2^T, && (3.19)\\ Z_1 &\succ 0, && (3.20)\\ P^c(Z_2) &= 0, && (3.21)\\ 0 &= Z_1 X + Z_2^T(I-A). && (3.22) \end{aligned}$$

Multiplying $(S - Z_1)$ to the LHS of (3.17) gives $Z_1 X = SX - I$. Substitute this with $X = (1/\alpha)(I-A)^T(I-A)$ to (3.22). This yields

$$\frac{1}{\alpha} S(I-A)^T(I-A) - I + Z_2^T(I-A) = 0.$$

Since $I - A$ is also invertible, we can write Z_2 in terms of A as

$$Z_2 = (I-A)^{-T} - \frac{1}{\alpha}(I-A)S. \tag{3.23}$$

It turns out that if there exists Z_2 and A such that $P(A) = 0$, $P^c(Z_2) = 0$ that make (3.23) satisfied then we can always construct X as an optimal solution. If (3.23) has many solutions, then the optimal X is not unique. Let n_A be the number of nonzero entries in A (the number of effective parameter), then the number of zero entries in Z_2 is n_A because $P^c(Z_2) = 0$. Then we can read from (3.23) that this is a set of n^2 nonlinear equations with n_A unknowns in A and $n^2 - n_A$ unknowns in Z_2. Out of n^2 equations, the n_A equations are set to zero on the LHS and the remaining $n^2 - n_A$ equations are set to the free entries of Z_2 on the LHS. Therefore, we can say that we try to solve n_A nonlinear equations with n_A unknowns of A. Determining the uniqueness of the solution cannot be obtained in general. One can apply the implicit function theorem that involves deriving the Jacobian matrix of

the nonlinear equation, which seems to be difficult for (3.23) as it is a nonlinear matrix equation. The condition (3.23) can be rewritten in matrix format including the constraints (3.18) and (3.21) as

$$0 = P^c \left[(I - P^c(A))^{-T} - \frac{1}{\alpha}(I - P^c(A))S \right], \tag{3.24}$$

by using $A = P(A) + P^c(A)$ and $Z_2 = P(Z_2) + P^c(Z_2)$. Moreover, even if we have found A satisfying (3.24), there is one more condition which is $Z_1 \succ 0$. This is equivalent to $X \succ S^{-1}$. In conclusion, optimal solutions exist under low rank assumption if there exists A satisfying

$$P^c \left[(I - P^c(A))^{-T} - \frac{1}{\alpha}(I - P^c(A))S \right] = 0, \tag{3.25}$$

$$\frac{1}{\alpha}(I - A)^T(I - A) \succ S^{-1}. \tag{3.26}$$

We consider that the above two conditions are complicated to conclude about the existence in general.

CHAPTER IV

SPARSE SEM WITH ℓ_1-NORM REGULARIZATION

In exploratory SEM analysis, one aims to discover a zero structure of A from the estimation process which reveals a causal structure of how one variable affects to another. An existing approach performs a local search starting from a null model (all path coefficients are zero) and sequentially allows the coefficient corresponding to the largest Lagrangian multiplier to be nonzero [8]. Another method is to start also from a null model and then add an extra path corresponding to the lowest minimized ML discrepancy selected among all possible paths. This scheme is referred to as *tree growth* as the model grows by a single entry in A at a time [10]. The most optimal but far from feasible approach is to perform a simple brute-force method (or known as *forest growth*) that searches through all possible patterns of zero structures in A with a fixed number of paths and chooses the model corresponding to the lowest minimized ML [10]. It is known that the number of all possible models grows exponentially to the number of variables (n), so it is not feasible as the problem dimension increases. Recently, [19] has proposed an overview of regularized structural equation modeling or RegSEM where ℓ_1-regularization is added to the cost objective of general SEM to promote the sparse relation structure between variables. The performance of RegSEM was evaluated from a typical example in the case that sample size varied. Their result showed that when sample size increased, the performance that was measured by averaged false positive percentage also increased as they expected.

In this chapter, we propose a convex formulation for exploratory SEM problem by applying a widely-used sparse optimization with ℓ_1-norm. The effectiveness of this approach has been well-understood and found applications in many fields including system identification [20], statistical learning [21, §6] or control [22] since the ℓ_1-norm penalty or *lasso* has been introduced [14, §3]. The convex formulation we propose is

$$\begin{aligned} \text{minimize} \quad & -\log\det X + \mathbf{tr}(SX) + 2\gamma \sum_{(i,j)\notin I_A} |A_{ij}|, \\ \text{subject to} \quad & \begin{bmatrix} X & (I-A)^T \\ I-A & \Psi \end{bmatrix} \succeq 0, \\ & 0 \preceq \Psi \preceq \alpha I, \quad P(A) = 0, \end{aligned} \tag{4.1}$$

with variables $X \in \mathbf{S}^n$, $A \in \mathbf{R}^{n\times n}$ and $\Psi \in \mathbf{S}^n$, and the dual of (4.1) is

$$\begin{aligned} \text{maximize} \quad & \log\det(S - Z_1) - 2\,\mathbf{tr}(Z_2) - \alpha\,\mathbf{tr}(Z_4) + n, \\ \text{subject to} \quad & \begin{bmatrix} Z_1 & Z_2^T \\ Z_2 & Z_4 \end{bmatrix} \succeq 0, \\ & |(Z_2)_{ij}| \le \gamma, \quad \forall\, (i,j) \notin I_A, \end{aligned} \tag{4.2}$$

with variable $Z = \begin{bmatrix} Z_1 & Z_2^T \\ Z_2 & Z_4 \end{bmatrix} \in \mathbf{S}^{2n}$. Derivation of the dual is explained in Appendix 9.3.

From (4.1), let

$$h(A) = \sum_{(i,j)\notin I_A} |A_{ij}| \tag{4.3}$$

and h is referred to as an ℓ_1-like penalty function (or regularization) as it resembles the 1-norm of a matrix except that only those entries belonging to I_A are penalized. The user gets to hypothesize about the *known* location of zeros in A which is encoded as the index set I_A. If the user has no prior knowledge about the zero locations in A at all then at least we apply the constraint $P(A) = \mathbf{diag}(A) = 0$ since there must be no path from one variable to itself. The constraint $P(A) = 0$ means we believe these locations A must be exactly zero. For $(i,j) \notin I_A$, then we are not uncertain if A_{ij} would be zero or not, so we enforce the ℓ_1 norm on these terms and let the regularization promote their sparsity which is controlled by the *regularization parameter* $\gamma > 0$. We refer the problem (4.1) as *sparse SEM*.

We see that the sparsity of the optimal path coefficient A can be controlled via the value of the penalty parameter, γ, *e.g.*, the larger γ, the sparser the matrix A is. In the Appendix 9.4, we will show that there exists a critical value of γ, denoted by $\gamma_{\max}$ in the sense that if

$$\gamma \geq \gamma_{\max}$$

then the optimal solution of A in (4.1) is the zero matrix. Moreover, the value of $\gamma_{\max}$ can be calculated in advance and depends on S (parameter of the problem). This means it is unnecessary to vary γ arbitrarily in the problem, and we can use $\gamma_{\max}$ as an upper bound of the range of γ used for varying the sparsity patterns of A, or for controlling the sparseness of A.

Another property of (4.1) is that its cost objective is not differentiable at $A = 0$ due to the term $|A_{ij}|$. As a result, KKT conditions for nonsmooth optimization problems are stated from the concept of subgradients and subgradient calculus which are provided in the Appendix 9.2. This provides a tool for the derivation of $\gamma_{\max}$. The nonsmooth property also makes solving a high-dimensional problem nontrivial since standard gradient-based methods cannot be applied.

Problem assumption. We will assume that S (the sample covariance matrix) is positive definite. Otherwise, the sparse SEM problem (4.1) is unbounded below. To show this, assume that S has the eigenvalue decomposition $S = UDU^T$. Then it follows that $\mathbf{tr}(SX) = \mathbf{tr}(UDU^TX) = \mathbf{tr}(DU^TXU)$. Let $Y = U^TXU$ and since $\det X = \det Y$, we can write $f(X,A) = f(Y,A) = -\log\det Y + \mathbf{tr}(DY) + 2\gamma\sum_{(i,j)\notin I_A}|A_{ij}|$. For minimizing $f(Y,A)$ with constraint in (4.1), it has at least a feasible point that yields an unbounded value in the cost function if S is merely positive semidefinite. For example, we can choose A to be a zero matrix, providing $\sum_{(i,j)\notin I_A}|A_{ij}| = 0$, but if $S \succeq 0$, then $d_{ii} = 0$ for some i, and one of the feasibility condition requires only $Y \succeq (1/\alpha)I$, so that we can choose Y to be diagonal where y_{ii} is chosen to be arbitrarily large. Hence, the term

$\mathbf{tr}(DY) = 0$ but $-\log\det Y \to -\infty$, leading the cost function to be unbounded below. We comment that the assumption on the positiveness of S might not be held in some applications when the number of variables are much bigger than the number of samples. In such case, if we replace S by $\tilde{S} = S + \epsilon I$ where $\epsilon > 0$, then the cost objective of (4.1) is replaced by

$$-\log\det X + \mathbf{tr}(SX) + \epsilon\|X\|_* + 2\gamma h(A)$$

where $\|X\|_* = \sum_{k=1}^{n} \sigma_k(X)$ or the nuclear norm of X. This follows from

$$\mathbf{tr}(\tilde{S}X) = \mathbf{tr}(SX) + \epsilon\,\mathbf{tr}(X).$$

Moreover, for a positive definite X, we have $\sigma(X) = \sqrt{\lambda(X^TX)} = \sqrt{\lambda(X^2)} = \sqrt{\lambda(X)^2} = \lambda(X)$. Hence, $\mathbf{tr}(X) = \sum_k \lambda_k(X) = \sum_k \sigma_k(X) = \|X\|_*$. The nuclear norm of a matrix is well-known to be a convex approximation for the matrix rank. The new problem with the replacement of $\tilde{S}$ can then be interpreted as an SEM problem with a regularization term on X that promotes $\mathbf{rank}(X)$ to be small. However, X cannot be low rank due to the implicit constraint in the $\log\det$ function that X must be invertible.

Solution pathway. Since we can derive the critical regularization parameter, $\gamma_{\max}$, such that for any $\gamma \geq \gamma_{\max}$, the solution of A_{ij} for $(i,j) \notin I_A$ reaches to zero, we plot the solution pathway by plotting the entries of A_{ij} for $(i,j) \notin I_A$ against varied γ as shown in Figure 4.1. This plot illustrates that as γ increases some of A_{ij} become zero and once an entry of A becomes zero for a value of γ then it stays zero as γ increases. When $\gamma \geq \gamma_{\max}$, all entries of A_{ij} subject to zero constraints become zero. In short, we obtain a sparser path matrix as we increase the penalty parameter.

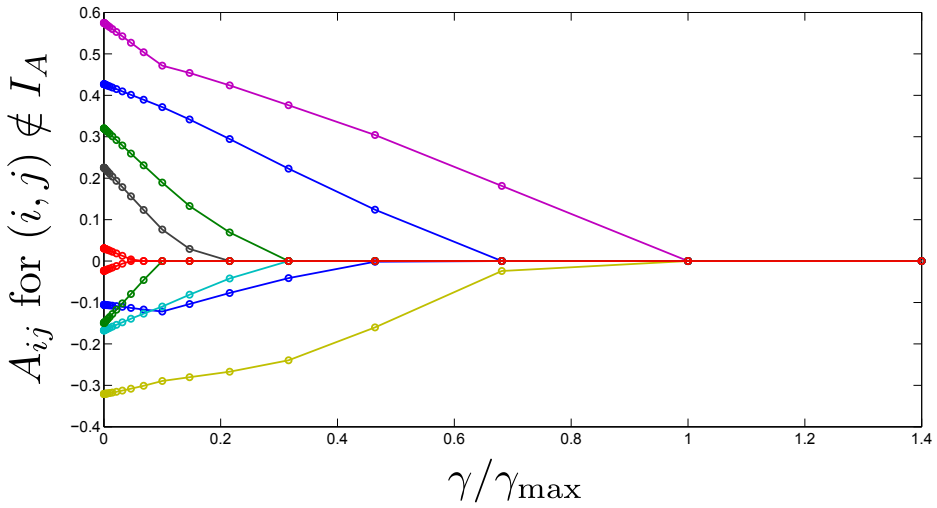


Figure 4.1: An example of sparse SEM implementation. When $\gamma = \gamma_{\max}$, our sparse SEM provides the sparsest solution, *i.e.*, all entries of A_{ij} for $(i,j) \notin I_A$ are set to zero.

4.1 KKT conditions

The KKT conditions are derived as the optimality condition for the optimal solution to (4.1). These conditions are:

Zero gradient of the Lagrangian

$$X = (S - Z_1)^{-1}. \tag{4.4}$$

Primal feasibility

$$(I - A)^T \Psi^{-1} (I - A) \preceq X, \tag{4.5}$$

$$0 \prec \Psi \preceq \alpha I, \tag{4.6}$$

$$P(A) = 0. \tag{4.7}$$

Dual feasibility

$$Z = \begin{bmatrix} Z_1 & Z_2^T \\ Z_2 & Z_4 \end{bmatrix} \succeq 0, \tag{4.8}$$

$$\|P^c(Z_2)\|_\infty \leq \gamma. \tag{4.9}$$

Complementary slackness condition

$$\mathbf{tr}\left(\begin{bmatrix} Z_1 & Z_2^T \\ Z_2 & Z_4 \end{bmatrix} \begin{bmatrix} X & (I-A)^T \\ I - A & \Psi \end{bmatrix} \right) = 0, \tag{4.10}$$

$$\mathbf{tr}\left(Z_4(\Psi - \alpha I)\right) = 0. \tag{4.11}$$

We will use these conditions to analyze the solution properties later throughout this thesis.

4.2 Low rank solutions of the sparse SEM

The solution of our sparse SEM is useful if $X = (I - A)^T \Psi^{-1} (I - A)$ at optimum as we can use X as the estimate of Σ^{-1}. This occurs if and only if the rank of

$$\begin{bmatrix} X & (I-A)^T \\ (I - A) & \Psi \end{bmatrix}$$

is n at optimum. To show this in detail, we follow the analysis explained in section 3.3. Referring to the complementary slackness condition (4.10)], we have

$$\underbrace{\begin{bmatrix} Z_1 & Z_2^T \\ Z_2 & Z_4 \end{bmatrix}}_{Z} \underbrace{\begin{bmatrix} X & (I-A)^T \\ I - A & \Psi \end{bmatrix}}_{W} = 0. \tag{4.12}$$

Similarly to our previous analysis in section 3.3, we obtain *low rank* solutions of Z and W when the optimal primal of (4.1) and the optimal dual of (4.2) satisfy

$$X = (I - A)^T \Psi^{-1} (I - A) \text{ or equivalently } \mathbf{rank}(Z) = n.$$

Furthermore, when this holds, $\mathbf{rank}(Z_4) = n$ and from (4.11), it gives $\Psi = \alpha I$, *i.e.*, the estimated covariance error becomes a diagonal matrix.

CHAPTER V

EXPLORATORY SEM

A model that explains the dynamics of a complex system typically has a great number of model parameters. The estimation of such models may encounter an overfitting problem and the model estimate suffers from having a large variance. Parsimonious models are then typically preferred and can be obtained by restricting some of the model parameters on some domain. In our case, putting restrictions on the zero pattern in the path matrix is in fact a method of obtaining a sparse model. As we see in chapter 4 that controlling the regularization parameter in the sparse SEM problem can provide path matrix solution with various sparsity patterns. If γ is large then the path matrix A contains many zeros, resulting in a simple interpretation on the estimated causal structure but the goodness of fit becomes bigger. Therefore, choosing an appropriate value of γ is a trade off between choosing the solution to explain a causal structure in a simple way and to best describe data in a certain level.

In this chapter, we explain a criterion for selecting a suitable choice of γ. In statistical approaches, there are several statistic criterions, for instance, Akaike Information Criterion (AIC) [23, 24], Akaike's Final Prediction-Error Criterion (FPE) [23] or Bayesian Information Criterion (BIC) [25, §7]. BIC is known to prefer a simpler model since the penalty on the model complexity is higher relatively to other criterions. Moreover, it can be shown that BIC chooses the correct model with probability reaching to one when the number of sample sizes grows to infinity. AIC is the first model selection criterion which has been wildly accepted and is known to perform poorly when the number of sample sizes is small compared with the number of effective parameters. To solve this problem, the corrected AIC or AICc was developed to improve the performance of AIC [26]. Kullback Information Criterion (KIC) is a recent model selection criterion based on Kullback's symmetric divergence [27]. Similarly to AICc, the corrected KIC (KICc) [28] was developed to reduce bias and improve model selection for a small-sample setting. For these reasons, we compare the performance of each model selection criterion, *i.e.*, BIC, AIC, AICc, KIC and KICc for SEM which are given by

$$\begin{aligned}
\mathrm{BIC} &= -2\mathcal{L} + d\log N, \\
\mathrm{AIC} &= -2\mathcal{L} + 2d, \\
\mathrm{AICc} &= -2\mathcal{L} + \frac{2dN}{N-d-1}, \\
\mathrm{KIC} &= -2\mathcal{L} + 3d, \\
\mathrm{KICc} &= -2\mathcal{L} + \frac{(d+1)(3N-d-2)}{N-d-2} + \frac{d}{N-d},
\end{aligned} \tag{5.1}$$

where d : the number of effective parameters of model, N is the number of samples and

$$\mathcal{L} = \frac{N}{2}\left(-\log\det\hat{\Sigma} - \mathbf{tr}(S\hat{\Sigma}^{-1})\right)$$

is the loglikelihood function of samples $Y_1, Y_2, \ldots, Y_N$. We note that this is a fair comparison between each criterion since each of them consists of two terms; the score from the first time represents the model fit and the score from the second term depends on the number of effective parameters.

To learn the best causal structure of path matrices, we can choose a range of γ and then solve sparse SEM (4.1) for each of those values, resulting in the estimated path matrices having different sparsity patterns ranging from dense to spareset. Each of the estimated sparsity pattern is then used as a sparsity constraint on A encoded in the primal convex SEM (3.1) and we solve for the optimal path matrix A equipped with a sparsity pattern, yielding a candidate model. We repeat this process using all the values of γ and obtain a set of candidate models and then compute all model selection criterion scores in (5.1) for each of them. The best model of each model selection criterion corresponds to the one with the minimum model selection scores. In short, we use the sparse SEM to select a finite number of sparsity patterns in A and use the primal convex SEM to provide the best estimate of the path matrix corresponding to the sparsity pattern selected from the model selection criterion score. This procedure is illustrated in Figure 5.1.

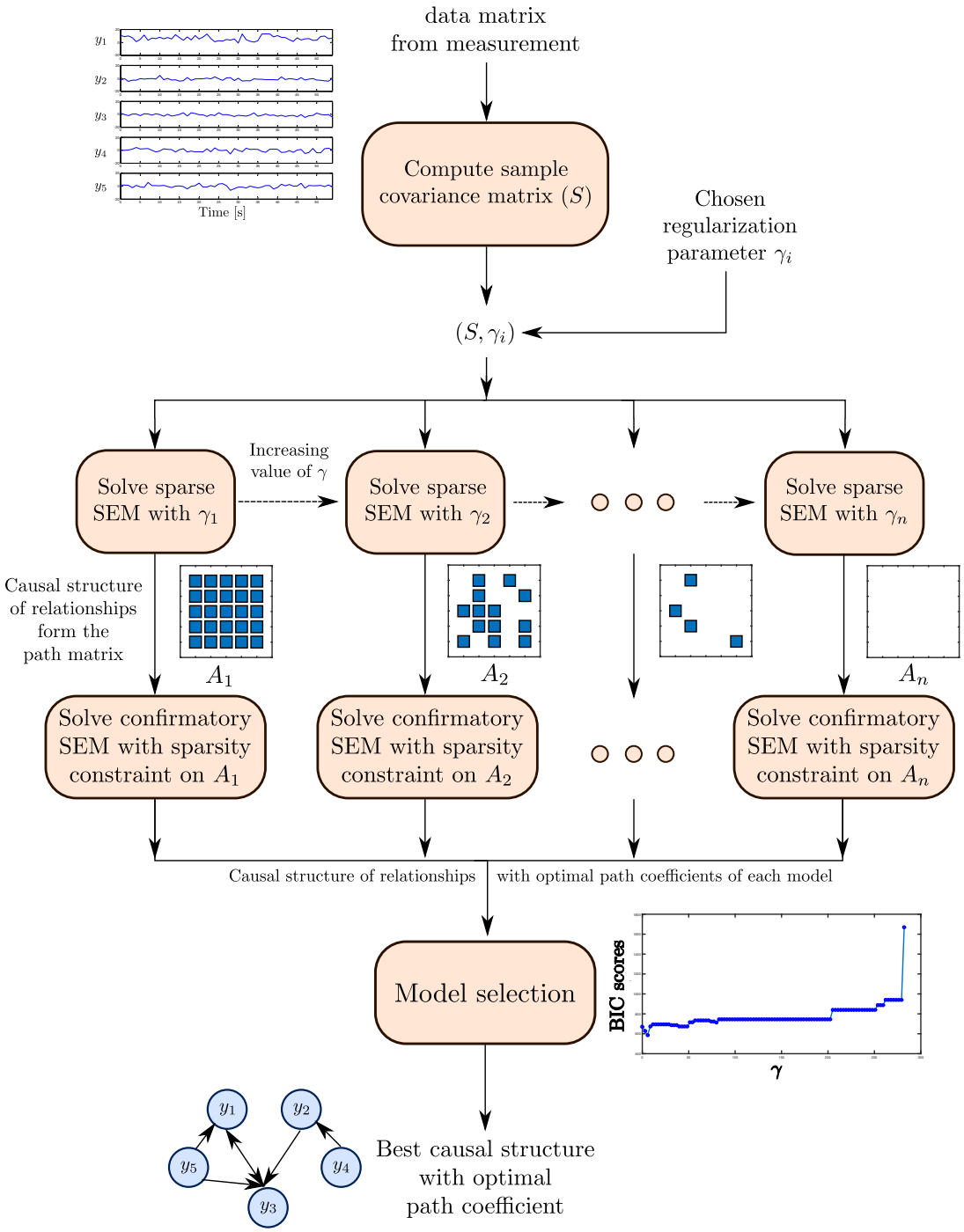


Figure 5.1: Procedure of learning a causal structure of path matrices.

CHAPTER VI

ALTERNATING DIRECTION METHOD OF MULTIPLIERS

In this chapter, we present efficient numerical methods for solving our two estimation formulations, the primal convex SEM and the sparse SEM. These methods mainly rely on a kind of proximal algorithm, which is called alternating direction method of multipliers or ADMM and proposed by [21]. ADMM is an algorithm that solves the general convex optimization problem by using the splitting technique to objective function and introducing some auxiliary variables, so that before using ADMM, we can rearrange the general problem

$$\begin{aligned} \text{minimize} \quad & f(x), \\ \text{subject to} \quad & x \in \mathcal{C}, \end{aligned}$$

where f is convex and $\mathcal{C}$ is a convex constraint set, into the form

$$\begin{aligned} \text{minimize} \quad & f(x) + g(z), \\ \text{subject to} \quad & Ax + Bz = c, \end{aligned} \tag{6.1}$$

with variables $x \in \mathbf{R}^n, z \in \mathbf{R}^m$ and given $A \in \mathbf{R}^{p\times n}, B \in \mathbf{R}^{p\times m}$, respectively. The function f and g are assumed to be closed proper convex. In each update step, the algorithm minimizes the augmented Lagrangian, defined by

$$L_\rho(x, z, y) = f(x) + g(z) + y^T(Ax + Bz - c) + \frac{\rho}{2}\|Ax + Bz - c\|_2^2,$$

respect to x and z, sequentially. In the above equation, ρ is called the *penalty parameter* in which its value relates to the speed of convergence and enforcing the equality constraint. The update rule of ADMM is described as follows.

$$\begin{aligned} x^{k+1} &= \operatorname*{argmin}_x L_\rho(x, z^k, y^k), \\ z^{k+1} &= \operatorname*{argmin}_z L_\rho(x^{k+1}, z, y^k), \\ y^{k+1} &= y^k + \rho(Ax^{k+1} + Bz^{k+1} - c). \end{aligned} \tag{6.2}$$

From (6.1), let us define $r^{k+1} = Ax^{k+1} + Bz^{k+1} - c$ as the primal residual at iteration $k+1$ and $s^{k+1} = \rho A^T B(z^{k+1} - z^k)$ as the dual residual at iteration $k+1$. The iterations (6.2) should reasonably stop when the primal and dual residual are less than some tolerance values as,

$$\|r^k\|_2 \le \epsilon^{\text{pri}} \quad \text{and} \quad \|s^k\|_2 \le \epsilon^{\text{dual}},$$

where ϵ^{pri} and ϵ^{dual} are the tolerances that depend on the dimension of problem and they can be computed by

$$\begin{aligned} \epsilon^{\text{pri}} &= \sqrt{p}\epsilon^{\text{abs}} + \epsilon^{\text{rel}} \max\{\|Ax^k\|_2, \|Bz^k\|_2, \|c\|_2\}, \\ \epsilon^{\text{dual}} &= \sqrt{n}\epsilon^{\text{abs}} + \epsilon^{\text{rel}}\|A^T y^k\|_2. \end{aligned}$$

ϵ^{abs} is an absolute tolerance depending on the scale of typical variable values and ϵ^{rel} is a relative tolerance that can be chosen by 10^{-3} or 10^{-4} depending on the application. The factors $\sqrt{p}$ and $\sqrt{n}$ account for the fact that ℓ_2 norms are in $\mathbf{R}^p$ and $\mathbf{R}^n$ respectively.

6.1 ADMM for solving primal convex SEM

In this section, we apply an ADMM algorithm to solve our primal conven SEM formulation (3.2). The detail is described as follows. From section 3.3, the low rank solution holds when $\alpha = \lambda_{\min}(S)$ and it provides $\Psi = \alpha I$ at the optimum, so that, in this case we prefer to solve (3.2) with a replacement of $0 \preceq X_4 \preceq \alpha I$ by $X_4 = \alpha I$. Then the primal convex SEM that needs to be solved is in a form

$$\begin{array}{ll} \text{minimize} & -\log\det X_1 + \mathbf{tr}(SX_1), \\ \text{subject to} & X = \begin{bmatrix} X_1 & X_2^T \\ X_2 & X_4 \end{bmatrix} \succeq 0, \\ & X_4 = \alpha I, \quad P(X_2) = I, \end{array} \tag{6.3}$$

with variable $X \in \mathbf{S}^{2n}$ where $X_1, X_4 \in \mathbf{S}^n$ and $X_2 \in \mathbf{R}^{n\times n}$. To rearrange (6.3) into ADMM format, let us define a function $f : \mathbf{S}^{2n} \to \mathbf{R}$ given by $f(X) = -\log\det(X_1) + \mathbf{tr}(SX_1)$, the function $g_1 : \mathbf{S}^{2n} \to \mathbf{R}$ and $g_2 : \mathbf{R}^{2n\times 2n} \to \mathbf{R}$ are indicator functions given by

$$g_1(U) = \begin{cases} 0, & U \succeq 0, \\ \infty, & \text{otherwise}, \end{cases} \qquad g_2(V) = \begin{cases} 0, & P(V_2) = I, V_4 = \alpha I, \\ \infty, & \text{otherwise}. \end{cases}$$

Then the problem (6.3) can then be rearranged into ADMM format as

$$\begin{array}{ll} \text{minimize} & f(X) + g_1(U) + g_2(V), \\ \text{subject to} & X - U = 0, \quad X - V = 0, \end{array}$$

with variables X, U and $V \in \mathbf{S}^{2n}$. The ADMM algorithm starts with forming the augmented Lagrangian defined by

$$\begin{aligned} L_\rho(X, U, V, Y_1, Y_2) = &-\log\det(X_1) + \mathbf{tr}(SX_1) + g_1(U) + g_2(V) \\ &+ \mathbf{tr}(Y_1^T(X - U)) + \mathbf{tr}(Y_2^T(X - V)) \\ &+ \frac{\rho}{2}\|X - U\|_F^2 + \frac{\rho}{2}\|X - V\|_F^2, \end{aligned}$$

where ρ is called the *penalty parameter* and the speed of convergence depends on the value of ρ. Let us denote X and X^+ the variables in current and next iteration. The update rule of ADMM is to minimize $L_\rho(X, U, V, Y_1, Y_2)$ with respect to X, U, V independently and can be described as follows.

$$\begin{aligned} X^+ = &\operatorname*{argmin}_X \; f(X) + \mathbf{tr}(Y_1^T(X - U)) + \mathbf{tr}(Y_2^T(X - V)) \\ &\qquad + \frac{\rho}{2}\|X - U\|_F^2 + \frac{\rho}{2}\|X - V\|_F^2, \\ U^+ = &\operatorname*{argmin}_U \; g_1(U) + \mathbf{tr}(Y_1^T(X - U)) + \frac{\rho}{2}\|X - U\|_F^2, \\ V^+ = &\operatorname*{argmin}_V \; g_2(V) + \mathbf{tr}(Y_2^T(X - V)) + \frac{\rho}{2}\|X - V\|_F^2, \end{aligned}$$

where we can show that the X, U and V updates can be derived into a closed form, so that we can compute these updates efficiently.

X-update. To find the X-update step, we minimize $L_\rho(X, U, V, Y_1, Y_2)$ with respect to X which is the problem

$$\underset{X}{\text{minimize}} - \log\det(X_1) + \mathbf{tr}(SX_1) + \rho\|X - M\|_F^2,$$

where $M = \frac{1}{2}(U + V) - \frac{1}{2\rho}(Y_1 + Y_2) \in \mathbf{S}^{2n}$. Suppose $M = \begin{bmatrix} M_1 & M_2^T \\ M_2 & M_4 \end{bmatrix}$. The zero gradient condition is

$$\nabla_X L_\rho(X, U, V, Y_1, Y_2) = \begin{bmatrix} -X_1^{-1} + S & 0 \\ 0 & 0 \end{bmatrix} + 2\rho(X - M) = 0, \tag{6.4}$$

with an implicit constraint from the domain of f that $X_1 \succ 0$. We can apply the method based on eigenvalue decomposition from [21, §6.5] to show that the zero gradient condition on the $(1, 1)$ block:

$$2\rho X_1 - X_1^{-1} = 2\rho M_1 - S, \tag{6.5}$$

can be achieved with a positive definite X_1. The detail starts with taking an eigenvalue decomposition on the RHS of (6.5), providing

$$2\rho X_1 - X_1^{-1} = Q\Lambda Q^T, \tag{6.6}$$

where Q is a matrix of eigenvector and $\Lambda = \mathbf{diag}(\lambda_1, \ldots, \lambda_n)$. Then we multiply Q^T on the left and Q on the right of (6.6). By the property that $Q^TQ = QQ^T = I$, now (6.6) becomes

$$2\rho\tilde{X}_1 - \tilde{X}_1^{-1} = \Lambda, \tag{6.7}$$

where $\tilde{X}_1 = Q^T X_1 Q$. The above equation leads us to find the positive number of $(\tilde{X}_1)_{ii}$ satisfying $2\rho(\tilde{X}_1)_{ii} - (\tilde{X}_1)_{ii}^{-1}) = \lambda_i$. By the quadratic formulation, the solution of (6.7),

$$(\tilde{X}_1)_{ii} = \frac{\lambda_i + \sqrt{\lambda_i^2 + 8\rho}}{4\rho}$$

is always positive. Therefore $X_1 = Q\tilde{X}_1 Q^T$. Other blocks of X are simply $X_2 = M_2$ and $X_4 = M_4$. The solution X_1, X_2 and X_4 satisfy the optimality condition (6.4) so these are the closed form solution in X-update. Hence

$$X^+ = \begin{bmatrix} Q\tilde{X}_1 Q^T & M_2^T \\ M_2 & M_4 \end{bmatrix}.$$

The main computational cost of this step depends on the eigenvalue decomposition of $2M_1 - S$ which is a symmetric matrix.

U-update. To find the U-update step, we minimize $L_\rho(X, U, V, Y_1, Y_2)$ with respect to U which is the problem

$$\underset{U}{\text{minimize}}\ g_1(U) + \mathbf{tr}(Y_1^T(X-U)) + \frac{\rho}{2}\|X-U\|_F^2,$$

and is equivalent to

$$\underset{U\succeq 0}{\text{minimize}} \quad \|U-M\|_F^2, \tag{6.8}$$

where $M = X + \frac{1}{\rho}Y_1$. This is a projection problem onto the positive definite cone and it has a closed form solution. The detail is explained in Appendix 9.6. Hence

$$U^+ = \Pi_{\mathcal{C}}(M) \text{ where } C = \mathbf{S}_+^{2n}.$$

V-update. To find the V-update step, we minimize $L_\rho(X, U, V, Y_1, Y_2)$ with respect to V which is the problem

$$\underset{V}{\text{minimize}}\ g_2(V) + \mathbf{tr}(Y_2^T(X-V)) + \frac{\rho}{2}\|X-V\|_F^2,$$

and is equivalent to

$$\begin{aligned} \underset{V}{\text{minimize}} \quad & \|V-M\|_F^2, \\ \text{subject to} \quad & P(V_2) = I, \quad V_4 = \alpha I, \end{aligned}$$

where $M = (X + \frac{1}{\rho}Y_2)$. From the two constraints in the problem, we can write a feasible V as

$$V = \begin{bmatrix} V_1 & V_2^T \\ V2 & V_4 \end{bmatrix} = \begin{bmatrix} V_1 & P(V_2^T) + P^c(V_2^T) \\ P(V_2) + P^c(V_2) & V_4 \end{bmatrix} = \begin{bmatrix} V_1 & I + P^c(V_2^T) \\ I + P^c(V_2) & \alpha I \end{bmatrix}.$$

Then, the cost function

$$\|V-M\|_F^2 = \|V_1 - M_1\|_F^2 + 2\|(I + P^c(V_2)) - M_2\|_F^2 + \|\alpha I - M_4\|_F^2,$$

is minimized by the optimal V given by

$$V = \begin{bmatrix} M_1 & P^c(M_2^T) + I \\ P^c(M_2) + I & \alpha I \end{bmatrix}.$$

Hence

$$V^+ = \begin{bmatrix} M_1 & P^c(M_2^T) + I \\ P^c(M_2) + I & \alpha I \end{bmatrix}.$$

The update rules of algorithm are described again as follows.

ADMM for solving primal convex SEM (6.3)**.** All variables in the algorithms consist of $X, U, V, Y_1, Y_2 \in \mathbf{S}^{2n}$. Initialize $\alpha = \lambda_{\min}(S)$, Y_1, Y_2, U as identity matrix and V such that $V_1 = 0, V_2 = I, V_4 = \alpha I$. Repeat these steps:

- Setting $M = \frac{1}{2}(U+V) - \frac{1}{2\rho}(Y_1+Y_2)$, where $M = \begin{bmatrix} M_1 & M_2^T \\ M_2 & M_4 \end{bmatrix}$, then

  $X^+ = \begin{bmatrix} Q\tilde{X}_1 Q^T & M_2^T \\ M_2 & M_4 \end{bmatrix}$,

  $U^+ = \Pi_{\mathcal{C}}(X - \frac{1}{\rho}Y_1)$ where $C = \mathbf{S}_+^{2n}$,

- Setting $M = X + \frac{1}{\rho}Y_2$, where $M = \begin{bmatrix} M_1 & M_2^T \\ M_2 & M_4 \end{bmatrix}$, then

  $V^+ = \begin{bmatrix} M_1 & P^c(M_2^T) + I \\ P^c(M_2) + I & \alpha I \end{bmatrix}$,

- $Y_1^+ = Y_1 + \rho(X^+ - U^+)$,
- $Y_2^+ = Y_2 + \rho(X^+ - V^+)$,

until primal residual (r) and dual residual (s) are less than some tolerance values:

$$\|r\|_F = \left\| \begin{bmatrix} X - U \\ X - V \end{bmatrix} \right\|_F \leq \epsilon^{\text{pri}}, \quad \|s\|_F = \rho \left\| \begin{bmatrix} X^+ - X \\ U^+ - U \\ V^+ - V \end{bmatrix} \right\|_F \leq \epsilon^{\text{dual}}.$$

The tolerance values ϵ^{pri} and ϵ^{dual} can be computed by

$$\epsilon^{\text{pri}} = 2n\epsilon^{\text{abs}} + \epsilon^{\text{rel}} \max\{\|X\|_F, \|U\|_F, \|V\|_F\},$$
$$\epsilon^{\text{dual}} = 2n\epsilon^{\text{abs}} + \epsilon^{\text{rel}} \max\{\|Y_1\|_F, \|Y_2\|_F\},$$

where $\epsilon^{\text{abs}} = 10^{-6}$ and $\epsilon^{\text{rel}} = 10^{-6}$ are chosen.

6.2 ADMM for solving sparse SEM with ℓ_1-norm regularization

In this section, we apply ADMM to solve our sparse SEM. The detail is explained as follows. Let us recall our sparse SEM (4.1) again. We can make a change of variables of this problem by letting

$$X = \begin{bmatrix} X_1 & X_2^T \\ X_2 & X_4 \end{bmatrix}, \quad X_4 = \Psi, \quad X_2 = I - Z, \quad Z = A.$$

Then, the problem (4.1) becomes

$$\begin{array}{ll} \underset{X,Z}{\text{minimize}} & -\log\det X_1 + \mathbf{tr}(SX_1) + 2\gamma \sum_{(i,j)\notin I_A} |Z_{ij}|, \\ \text{subject to} & X = \begin{bmatrix} X_1 & X_2^T \\ X_2 & X_4 \end{bmatrix} \succeq 0, \\ & 0 \preceq X_4 \preceq \alpha I, \\ & X_2 = I - Z, \\ & P(Z) = 0, \end{array} \tag{6.9}$$

with variables $X \in \mathbf{S}^{2n}$ and $Z \in \mathbf{R}^{n\times n}$. To rearrange (6.9) into ADMM format, let us define a function $f : \mathbf{S}^{2n} \to \mathbf{R}$ given by $f(X) = -\log\det(X_1) + \mathbf{tr}(SX_1)$, the function $g_1 : \mathbf{R}^{n\times n} \to \mathbf{R}$ given by $g_1(Z) = \sum_{(i,j)\notin I_A} |Z_{ij}|$, the function $g_2 : \mathbf{S}^{2n} \to \mathbf{R}$ and $g_3 : \mathbf{R}^{2n\times 2n} \to \mathbf{R}$ are indicator functions given by

$$g_2(U) = \begin{cases} 0, & U \succeq 0, \\ \infty, & \text{otherwise}, \end{cases} \qquad g_3(V) = \begin{cases} 0, & 0 \preceq V_4 \preceq \alpha I, \\ \infty, & \text{otherwise}. \end{cases}$$

Then the problem (6.9) becomes

$$\begin{array}{ll} \text{minimize} & f(X) + 2\gamma g_1(Z), \\ \text{subject to} & X_2 = I - Z, \quad P(Z) = 0, \\ & X - U = 0, \quad X - V = 0, \end{array}$$

with variables $X, U, V \in \mathbf{S}^{2n}$ and $Z \in \mathbf{R}^{n\times n}$. The ADMM algorithm starts with forming the augmented Lagrangian defined by

$$\begin{aligned} L_\rho(X, Z, U, V, Y_1, Y_2, Y_3) =& f(X) + 2\gamma g_1(Z) + g_2(U) + g_3(V) \\ & + 2\,\mathbf{tr}(Y_1^T(X_2 + Z - I)) + \mathbf{tr}(Y_2^T(X - U)) + \mathbf{tr}(Y_3^T(X - V)) \\ & + \frac{\rho}{2}\|X_2 + Z - I\|_F^2 + \frac{\rho}{2}\|X - U\|_F^2 + \frac{\rho}{2}\|X - V\|_F^2. \end{aligned}$$

Let us denote X and X^+ the variables in current and next iteration. The update rule of ADMM is to minimize $L_\rho(X, Z, U, V, Y_1, Y_2, Y_3)$ with respect to X, Z, U and V, independently and can be

describes as follows

$$\begin{aligned}
X^+ = \quad & \underset{X}{\operatorname{argmin}} \; -\log\det(X_1) + 2\,\mathbf{tr}(Y_1^T(X_2+Z-I)) + \mathbf{tr}(Y_2^T(X-U)) + \mathbf{tr}(Y_3^T(X-V)) \\
& \qquad + \frac{\rho}{2}\|X_2+Z-I\|_F^2 + \frac{\rho}{2}\|X-U\|_F^2 + \frac{\rho}{2}\|X-V\|_F^2, \\
Z^+ = \quad & \underset{Z}{\operatorname{argmin}}\; g_1(Z) + \mathbf{tr}(Y_1^T(X_2+Z-I)) + \frac{\rho}{2}\|X_2+Z-I\|_F^2, \\
U^+ = \quad & \underset{U}{\operatorname{argmin}}\; g_2(U) + \mathbf{tr}(Y_1^T(X-U)) + \frac{\rho}{2}\|X-U\|_F^2, \\
V^+ = \quad & \underset{V}{\operatorname{argmin}}\; g_3(V) + \mathbf{tr}(Y_2^T(X-V)) + \frac{\rho}{2}\|X-V\|_F^2,
\end{aligned}$$

where we can show that the X, Z, U and V updates can be derived into a closed form.

X-update. For the X-update step, we minimize $L_\rho(X, Z, U, V, Y_1, Y_2, Y_3)$ with respect to X which is the problem

$$\underset{X}{\text{minimize}} \; -\log\det(X_1) + \mathbf{tr}(SX_1) + \frac{\rho}{2}\|X_2 - H\|_F^2 + \rho\|X - M\|_F^2,$$

where $H = I - Z - \frac{1}{\rho}Y_1 \in \mathbf{R}^{n\times n}$ and $M = \frac{1}{2}(U+V) - \frac{1}{2\rho}(Y_2+Y_3) \in \mathbf{S}^{2n}$. Suppose $M = \begin{bmatrix} M_1 & M_2^T \\ M_2 & M_4 \end{bmatrix}$. The zero gradient condition is

$$\begin{aligned}
&\nabla_X L_\rho(X, Z, U, V, Y_1, Y_2, Y_3) = \\
&\begin{bmatrix} -X_1^{-1} + S & 0 \\ 0 & 0 \end{bmatrix} + \rho \begin{bmatrix} 0 & (X_2 - H)^T \\ X_2 - H & 0 \end{bmatrix} + 2\rho \begin{bmatrix} X_1 - M_1 & 2(X_2 - M_2)^T \\ 2(X_2 - M_2) & X_4 - M_4 \end{bmatrix} = 0,
\end{aligned} \tag{6.10}$$

with an implicit constraint from domain of f that $X_1 \succ 0$. We can apply the method based on eigenvalue decomposition from [21, §6.5] to show that the zero gradient condition on the $(1,1)$ block:

$$2\rho X_1 - X_1^{-1} = 2\rho M_1 - S$$

can be achieved with a positive definite X_1. The detail of finding X_1 is same as the analysis explained in section 6.1 which we have shown that X_1 can be computed from $X_1 = Q\tilde{X}_1 Q^T$ where

$$(\tilde{X}_1)_{ii} = \frac{\lambda_i + \sqrt{\lambda_i^2 + 8\rho}}{4\rho},$$

and Q is a matrix of eigenvector from eigenvalue decomposition of $2\rho M_1 - S$. Other blocks of X are simply, given by $X_2 = (1/5)(H + 4M_2), X_4 = M_4$. The solution X_1, X_2 and X_4 satisfy the optimality condition (6.10) so these are the closed form solution in X-update. Hence

$$X^+ = \begin{bmatrix} Q\tilde{X}_1 Q^T & \frac{1}{5}(H + 4M_2)^T \\ \frac{1}{5}(H + 4M_2) & M_4 \end{bmatrix}.$$

The main computational cost of this step depends on the eigenvalue decomposition of a symmetric matrix.

Z-update. For the Z-update step, we solve the optimization problem as

$$\begin{array}{ll}\underset{Z}{\text{minimize}} & 2\gamma \sum_{(i,j)\notin I_A} |Z_{ij}| + \frac{\rho}{2}\|Z - M\|_F^2, \\ \text{subject to} & P(Z) = 0,\end{array}$$

where $M = (I - X_2 - \frac{1}{\rho}Y_1)$. The solution of minimizing the above problem can be performed by elementwise soft thresholding, given by

$$Z_{ij} = S_{2\gamma/\rho}(M_{ij})$$

where $S_k(a)$ is called *soft thresholding* operator [21], defined by

$$S_k(a) = \begin{cases} a-k, & a > k, \\ 0, & |a| \le k, \\ a+k, & a < -k, \end{cases} \tag{6.11}$$

or equivalently $S_k(a) = (a-k)_+ - (-a-k)_+$. Hence

$$Z_{ij}^+ = S_{2\gamma/\rho}(M_{ij}) \text{ for } (i,j) \notin I_A, \quad \text{and } Z_{ij}^+ = 0 \text{ for } (i,j) \in I_A.$$

U-update. For the U-update step, we solve the optimization problem as

$$\underset{U}{\text{minimize}}\ g_2(U) + \mathbf{tr}(Y_2^T(X-U)) + \frac{\rho}{2}\|X-U\|_F^2,$$

and is equivalent to

$$\underset{U\succeq 0}{\text{minimize}} \quad \frac{\rho}{2}\|U-M\|_F^2, \tag{6.12}$$

where $M = X + \frac{1}{\rho}Y_2$. This can be considered as a projection problem onto the positive definite cone and the closed form solution is explained in Appendix 9.6. Hence

$$U^+ = \Pi_{\mathcal{C}}(M) \text{ where } C = \mathbf{S}_+^{2n}.$$

V-update. For the V-update step, we solve the optimization as

$$\underset{V}{\text{minimize}}\ g_2(V) + \mathbf{tr}(Y_2^T(X-V)) + \frac{\rho}{2}\|X-V\|_F^2,$$

and is equivalent to

$$\begin{array}{ll}\underset{V}{\text{minimize}} & \frac{\rho}{2}\|V-M\|_F^2, \\ \text{subject to} & 0 \preceq V_4 \preceq \alpha I,\end{array} \tag{6.13}$$

where $M = X + \frac{1}{\rho}Y_3$. From the cost function

$$\|V-M\|_F^2 = \|V_1 - M_1\| + 2\|V_2 - M_2\| + \|V_4 - M_4\|,$$

The optimal V that minimizes this cost function is $V_1 = M_1$, $V_2 = M_2$, but for V_4 , it has a constraint such that $0 \preceq V_4 \preceq \alpha I$. To find the optimal V_4, it can be considered as a projection problem onto positive definite cone with an upper bound, *i.e.*,

$$\underset{V}{\text{minimize}} \|V_4 - M_4\| \quad \text{subject to} \quad 0 \preceq V_4 \preceq \alpha I.$$

The closed form solution of this problem is explained in Appendix 9.6 and if we define this solution as $V_4 = \tilde{M}_4$, hence

$$V = \begin{bmatrix} M_1 & M_2^T \\ M_2 & \tilde{M}_4 \end{bmatrix}.$$

Therefore

$$V^+ = \begin{bmatrix} M_1 & M_2^T \\ M_2 & \tilde{M}_4 \end{bmatrix}.$$

The update rules of algorithm are described again as follows.

ADMM for solving sparse SEM (6.9)**.** All variables in the algorithms consist of $X, U, V, Y_2, Y_3 \in \mathbf{S}^{2n}$ and $Z, Y_1 \in \mathbf{R}^{n\times n}$. Initialize $\alpha = \lambda_{\min}(S), Y_1, Y_2, Y_3$ and U as identity matrix, $Z = 0$, and V such that $V_1 = 0, V_2 = I, V_4 = \alpha I$. Repeat these steps:

- Setting $M = \frac{1}{2}(U + V) - \frac{1}{2\rho}(Y_2 + Y_3) \in \mathbf{S}^{2n}$ where $M = \begin{bmatrix} M_1 & M_2^T \\ M_2 & M_4 \end{bmatrix}$, and

  $H = I - Z - \frac{1}{\rho}Y_1 \in \mathbf{R}^{n\times n}$ then,

  $$X^+ = \begin{bmatrix} Q\tilde{X}_1 Q^T & \frac{1}{5}(H + 4M_2)^T \\ \frac{1}{5}(H + 4M_2) & M_4 \end{bmatrix},$$

- Setting $M = I - X_2 - \frac{1}{\rho}Y_1 \in \mathbf{R}^{n\times n}$ then, $Z_{ij}^+ = S_{2\gamma/\rho}(M_{ij}),\ \forall (i,j) \notin I_A, Z_{ij}^+ = 0,\ \forall (i,j) \in I_A,$
- Setting $M = X + \frac{1}{\rho}Y_2 \in \mathbf{R}^{2n\times 2n}$ then, $U^+ = \Pi_{\mathcal{C}}(M)$ where $C = \mathbf{S}_+^{2n}$,
- Setting $M = X + \frac{1}{\rho}Y_3 \in \mathbf{R}^{2n\times 2n}$ then, $V^+ = \begin{bmatrix} M_1 & M_2^T \\ M_2 & \tilde{M}_4 \end{bmatrix}$,
- $Y_1^+ = Y_1 + \rho(X_2^+ + Z^+ - I), \quad Y_2^+ = Y_2 + \rho(X^+ - U^+), \quad Y_3^+ = Y_3 + \rho(X^+ - V^+),$

until primal residual (r) and dual residual (s) are less than some tolerance values:

$$\|r_1\|_F = \left\| \begin{bmatrix} X - U \\ X - V \end{bmatrix} \right\|_F \le \epsilon^{\text{pri}}, \qquad \|s_1\|_F = \rho \left\| \begin{bmatrix} X^+ - X \\ U^+ - U \\ V^+ - V \end{bmatrix} \right\|_F \le \epsilon^{\text{dual}},$$

$$\|r_2\|_F = \|X_2 + Z - I\|_F \le \epsilon^{\text{pri}}, \quad \|s_2\|_F = \rho\|Z^+ - Z\|_F \le \epsilon^{\text{dual}}.$$

The tolerance values ϵ^{pri} and ϵ^{dual} can be computed by

$$\epsilon^{\text{pri}} = 2n\epsilon^{\text{abs}} + \epsilon^{\text{rel}} \max\{\|X\|_F, \|Z\|_F, \|U\|_F, \|V\|_F\},$$

$$\epsilon^{\text{dual}} = 2n\epsilon^{\text{abs}} + \epsilon^{\text{rel}} \max\{\|Y_1\|_F, \|Y_2\|_F, \|Y_3\|_F\},$$

where $\epsilon^{\text{abs}} = 10^{-6}$ and $\epsilon^{\text{rel}} = 10^{-6}$ are chosen.

CHAPTER VII

NUMERICAL RESULTS

All numerical experiments and corresponding results are demonstrated in this chapter. The main results are :

- Our primal convex SEM provides a low rank solution, leading to a useful solution for the original problem under a mild condition to α. This parameter can be chosen not to be arbitrarily large according to our guideline on its critical value. The estimation results of the formulation show that our solution can coincide with the original solution when the noise is homoskedastic.
- The sparse SEM can provide a good estimate of zero pattern in the path coefficient by using an appropriate regularization parameter . This choice typically depends on the number of samples, and the assumption of the true model.
- Our two formulations are applied in learning a brain network from fMRI data set and we found that the causal relation practically agreed with the findings from the original papers used this data set.

7.1 Results of the primal convex SEM formulation

This section provides the results that verify two main properties of the primal convex SEM formulation. One is that under an experimental condition on α, the solution is found to be low rank, making it useful as an estimate of the covariance matrix in the original SEM problem. Another property is on the magnitude of α that when it is large then the solution becomes zero and as a result, becomes meaningless. The last experiment in this section illustrate the estimation result when a true model is generated and we examine the estimation error under several values of the noise variances.

7.1.1 Low rank solutions

To find the condition that leads us to obtain a low rank solution, we generate S and α, and then solve the primal convex SEM and observe what condition on S and α provides the low rank solution. The simulation process is explained as follows. A sample covariance matrix (S) was generated as a positive definite (pdf) matrix in which its eigenvalues were controlled in the interval $[1, 20]$. Since A has special structures, *i.e.*, some entries of A are zero including diagonal entries, the structure of A was randomly generated by setting the sparsity of A about 30%. To vary α, we set $\alpha \in [0.5\lambda_{\min}(S), 5\lambda_{\min}(S)]$ with step size of 0.02. Using the problem parameters: S, α and sparsity

pattern of A, we then solve the primal convex SEM (3.1) for each α. To see a relationship between the low rank solution and $\mathbf{rank}(Z)$, we solve the dual of primal convex SEM (3.3) concurrently. Solving these two problems has been done by CVX package in MATLAB [17]. To compute $\mathbf{rank}(Z)$ for each α, we check the number of eigenvalues of Z that whose magnitude is smaller than a threshold value of 10^{-6}.

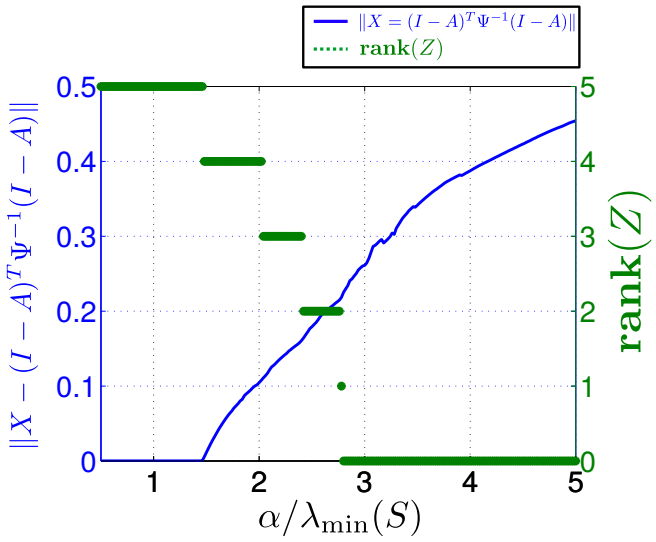


Figure 7.1: The rank of dual solution and the error of primal solution to a low rank solution versus α. The experiment was set up with $n = 5$ and varying $\alpha \in [0.5\lambda_{\min}(S), 5\lambda_{\min}(S)]$. The simulation results show that a low rank solution is obtained when $\mathbf{rank}(Z) = n = 5$ and when α is small enough relatively to $\lambda_{\min}(S)$.

Figure 7.1 illustrates the relationship between the low rank solution (rank of W in section 3.3) and rank of Z as α varies. We see that $X = (I-A)^T\Psi^{-1}(I-A)$ when $\mathbf{rank}(Z) = n$, and when α increases, X tends to be strictly greater than $(I-A)^T\Psi^{-1}(I-A)$ (so that $\mathbf{rank}(W) > n$) and therefore $\mathbf{rank}(Z)$ is decreasing. Figure 7.2 shows the difference between X (supposed to be the estimated Σ^{-1}) and $(I-A)^T\Psi^{-1}(I-A)$ using 50 runs of S with the same n, *i.e.*, solving the primal convex SEM with one sample of S produces a line in the figure. The norm of error, $\|X-(I-A)^T\Psi^{-1}(I-A)\|$, is zero when we obtain the low rank solution. We notice that the range of α resulting in the low rank solutions does not depend on n. This often occurs when $\alpha \leq \lambda_{\min}(S)$. Therefore, if we solve the primal convex SEM (3.1) instead of the original problem (2.6), we can heuristically choose $\alpha = \lambda_{\min}(S)$ to obtain a low rank solution.

7.1.2 Large value of α

In this section we show the result of **Theorem** 2. Let $\alpha_c = n/\mathbf{tr}(S^{-1})$ (the harmonic mean of eigenvalue of S). If $\alpha \leq \alpha_c$,then $Z = 0$ cannot be an optimal solution for dual problem. The experiment is setup with $n = 5$ and varying $\alpha \in [0.5\lambda_{\min}(S), 5\lambda_{\min}(S)]$, but in this experiment we generate each S as a positive definite matrix having the same α_c (the harmonic mean of eigenvalues of

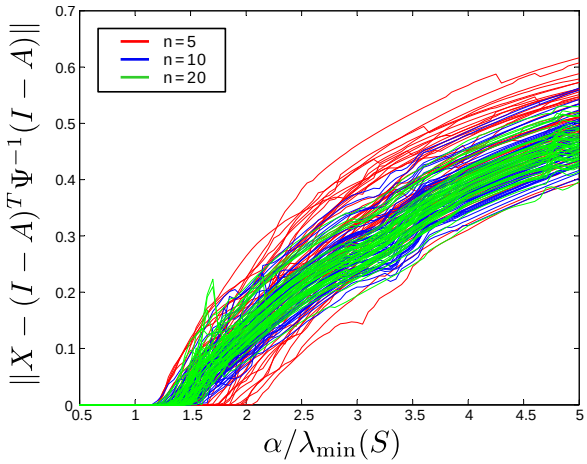


Figure 7.2: The error of primal solution to a low rank solution versus α. The experiment was set up by using $n = 5, 10, 20$ and varying $\alpha \in [0.5\lambda_{\min}(S), 5\lambda_{\min}(S)]$. Lines with the same color correspond to the result from using the same n. Each line in the same color is distinguished by each sample of S. The error between X and $(I-A)^T\Psi^{-1}(I-A)$ increases as α increases and is zero when α is sufficiently small relatively to the minimum eigenvalue of S.

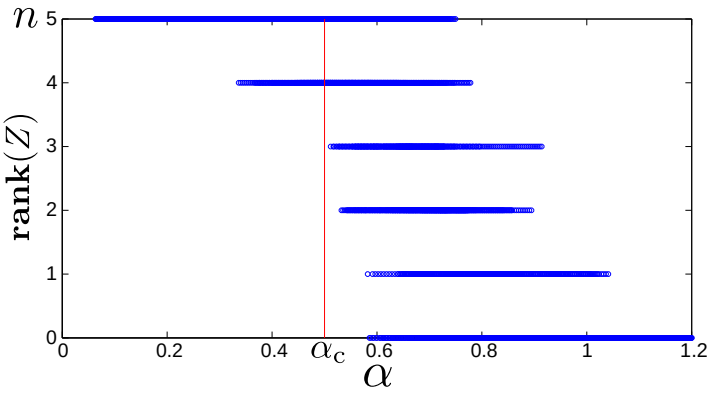


Figure 7.3: $\mathbf{rank}(Z)$ as α varies. The critical value α_c in the plot is the harmonic mean of eigenvalue of S, $\alpha_c = n/\,\mathbf{tr}(S^{-1})$. For each S, the condition $\mathbf{rank}(Z) = 0$ lies on RHS of α_c, meaning that if $Z = 0$, $\alpha > \alpha_c$.

S) to be 0.5. We then solve the dual of primal convex SEM and plot a relationship between $\mathbf{rank}(Z)$ and α.

From Figure 7.3, the experiment has been done with 50 samples of S and the result illustrates that for $\alpha \leq \alpha_c, Z$ cannot be zero. This plot can provide other information, for instance, $Z = 0$, when α is large enough, the portion that $\mathbf{rank}(Z) = n$ is approximately 74% and the portion that $\mathbf{rank}(Z) < n$ is approximately 36%, computed from 50 samples of S. Although, we cannot guarantee the relationship between the low rank solution and $\mathbf{rank}(Z)$ with α_c but this result can guide us that if we choose $\alpha < \alpha_c$, we have more chances to get the condition $\mathbf{rank}(Z) = n$ (or more chances to get a low rank solution).

7.1.3 Estimation results

In this section, we verify that if we suppose to know about the true path matrix, denoted by A_{true}, and variance of noise, σ^2, our estimation formulation can provide that the estimate is equal, $\hat{A} = A_{\text{true}}$. In this experiment, we firstly generate A_{true} with $n = 5$ and sparsity about 50% corresponding to 0 degree of freedom. We then generate noise covariance $\Psi = \sigma^2 I$ and we suppose to generate S by $S = \sigma^2(I - A_{\text{true}})^{-1}(I - A_{\text{true}})^{-T}$. For our approach, the result of simulation is illustrated in Figure 7.4. This simulation has been done by setting $n = 5$ as α varies in range $[0.0001, 0.02]$, using the step size of 0.0001.

In this plot, Figure 7.4 (top) shows the value of $\|A_{\text{true}} - \hat{A}\|$ as α varies. We observe that $\hat{A}$ reaches to A_{true} when α reaches to σ^2, meaning that our approach can provide $\hat{A}$ which is equal to A_{true} if we choose $\alpha = \sigma^2$. Figure 7.4 (middle and bottom) shows the result of perfect fitting, $X = S^{-1}$ and the value of objective of (3.1) is zero ($p^* = 0$). This result illustrates that we can get perfect fitting when α reaches to σ^2. But in real applications, we do not have information about noise variance, therefore we opt to choose $\alpha = \lambda_{\min}(S)$ so that it is guaranteed to obtain a low rank solution. From this choice of α, the value of estimated A is not significantly different from A_{true} and X reaches to S^{-1} but it is not exactly equal.

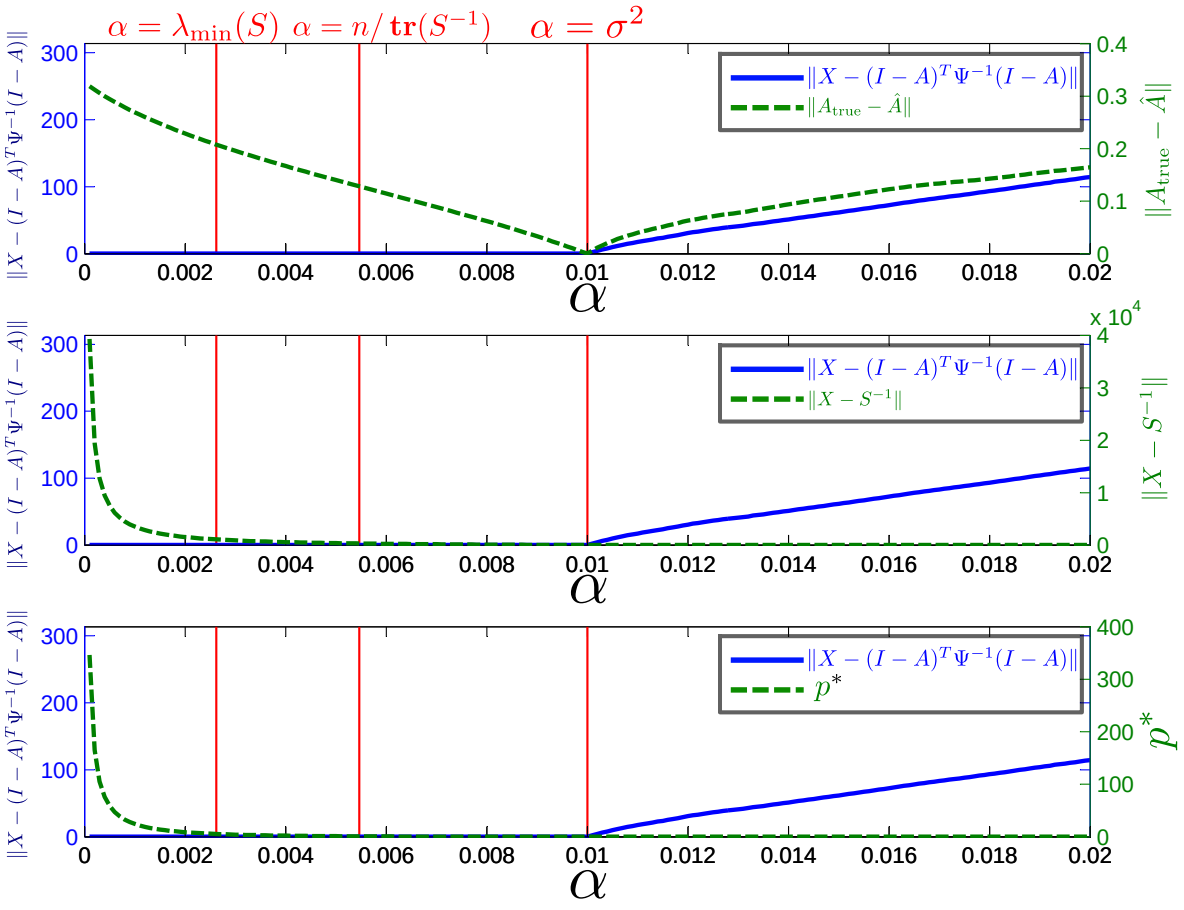


Figure 7.4: Simulation of estimation results with $n = 5$ as α varies. When $\alpha = \sigma^2$, we can get a low rank solution and a perfect fitting.

7.2 Results of sparse SEM with ℓ_1-norm regularization

This section illustrates the effectiveness of our sparse SEM formulation proposed in Section 4. The goal is to show that adding ℓ_1 penalty term could reveal the zero structure in the path matrix under some assumptions. To examine the performance of exploring the zero structures in the estimated path matrix, we define *positives* as nonzero entries and the *negatives* as zero entries of a matrix. Let $A_{\rm true}$ and $\hat{A}$ be the true and estimated path matrices, respectively. The four measures: TP, TN, FP and FN are described by

Measures	Definitions
TP (true positives)	number of nonzero entries in $\hat{A}$ and in $A_{\rm true}$ (correctly estimated nonzeros)
FP (false positives)	number of nonzero entries in $\hat{A}$ but not in $A_{\rm true}$ (falsely estimated nonzeros)
TN (true negatives)	number of zero entries in $\hat{A}$ and in $A_{\rm true}$ (correctly estimated zeros)
FN (false negatives)	number of zero entries in $\hat{A}$ but not in $A_{\rm true}$ (falsely estimated zeros)

As previously mentioned, FP and FN are considered to be two types of error so the total error and accuracy are defined as

$$\begin{aligned} \text{total error} &= (\text{FP} + \text{FN})/\text{number of estimated parameters}, \\ \text{accuracy} &= 1 - \text{total error}. \end{aligned}$$

Another way to evaluate the performance of learning a zero pattern is to compute TP and FP rates defined by

$$\begin{aligned} \text{TPR (TP rate)} &= \text{TP}/(\text{TP} + \text{FN}), \\ \text{FPR (FP rate)} &= \text{TN}/(\text{FP} + \text{TN}). \end{aligned}$$

We can plot TP rate against FP rate and this plot is commonly known as a receiver operating characteristic or ROC curve [29][§19.7]. From (4.1), ROC curve is obtained by varying the regularization parameter, γ, from 0 to its critical value, $\gamma_{\max}$. When $\gamma = 0$, our $\hat{A}$ is typically dense and we expect to see a high TP rate and a high FP rate. As γ increases, our sparse SEM provides a sparser solution and therefore we expect to see a decrease in FP rate. Finally a performance of our sparse SEM can be concluded via a pattern of ROC curve, saying that, a good performance should be reflected as an ROC curve above the diagonal line and lying towards the top left of the corner, meaning that, we can have a value of regularization parameter that yields a high TP rate and a low FP rate simultaneously at that point.

The first experiment of this section is to observe the effect of percentage of known number of zeros (relative to number of zeros), called degree of freedom (df), $\mathrm{df} =$ the number of known parameters $-$ the number of estimated parameters . The number of known parameters is $n(n-1)/2$, and, the number of estimated parameters is the sum of the number of entries in A and the number of entries in Ψ.

In this experiment, we use $n = 20$ and randomly generate a true path matrix A_{true} with a sparse density of 10% and generate a sample covariance matrix $S = (I - A_{\text{true}})^{-1}\Psi(I - A_{\text{true}})^{-T}$ by setting the noise covariance $\Psi = 0.1I$. We then solve (4.1) by assuming that the location of zeros in A_{true} is known in the amount of $0\%, 20\%, 50\%, 65\%$ and 80% of all zeros. The result is the ROC curve shown in Figure 7.5. One line of this plot is obtained by varying γ from 0 to $\gamma_{\max}$ and it is averaged from 25 samples. This result illustrates that if we do not have any assumptions on zero structure in A_{true}, *i.e.*, the constraint $P(A) = \mathbf{diag}(A)$, our accuracy is still more than 50% as this ROC curve stays above the diagonal line, and if we have more assumptions on zero structure in A_{true}, the accuracy of our learning method is improved. We note that if no known location of zeros is assumed (df is negative) then the problem (4.1) may not have a unique solution for a given γ, then the estimated zero structure of $\hat{A}$ may not be the same as the true matrix. However, if we assume more location of zeros (df is zero or positive), then the problem (4.1) could have a unique solution, and there exists a value of γ that yield a satisfactorily accurate zero structure in $\hat{A}$.

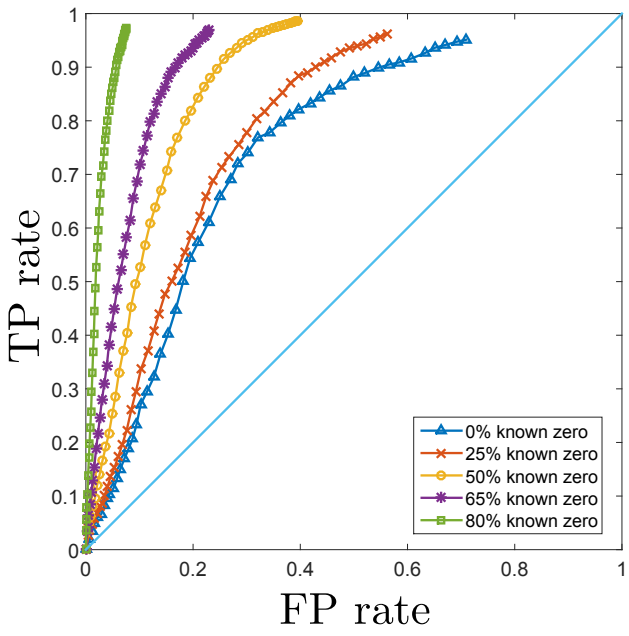


Figure 7.5: ROC curve as we vary regularization parameter γ. Knowing more correct zero structure in A_{true} provides the better accuracy of our learning causal structure method.

The second experiment is to observe the effect of the number of observations if the data generating process starts with $Y = AY + e$, where e denotes the noise from measurement. In this experiment, we assume that the percentage of known number of zeros about 50% or $\mathrm{df} = 0$. Firstly, we use $n = 20$ and randomly generate A_{true} with sparse density of 10% and generate the measurements Y from $Y = (I - A_{\text{true}})^{-1}e$ where e is normally distributed with variance of 0.1. We then compute a sample covariance matrix from measurements Y. With the same A_{true}, we vary the number of observation to $100, 1000$ and 5000 respectively. The result is the ROC curve shown in Figure 7.6. One line of this plot is obtained by varying γ from 0 to $\gamma_{\max}$ and it is averaged from 25 samples of S.

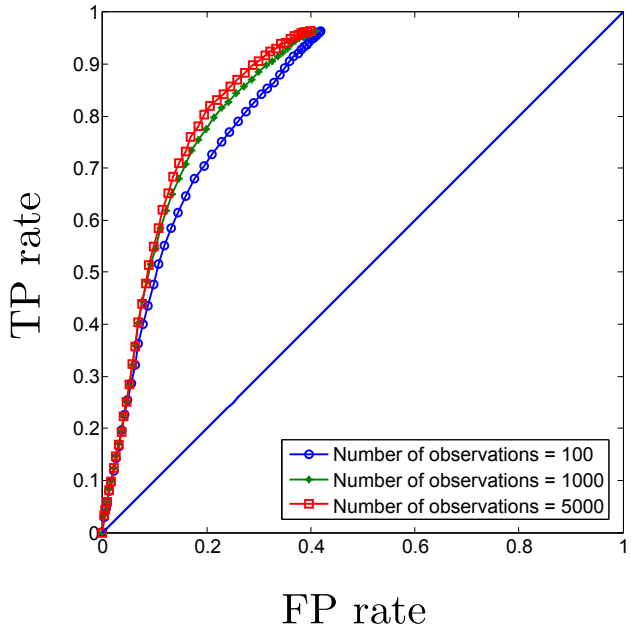


Figure 7.6: ROC curve as we vary regularization parameter γ. When the number of observation increases, the performance of our exploratory SEM also significantly increases.

This result illustrates that when the number of observation increases, the performance of our sparse SEM also significantly increases.

7.3 Results of exploratory SEM

In this section, we illustrate the effectiveness of exploring the relationships among variables by applying the two proposed formulations in the scheme explained in Figure 5.1. The goal of this experiment is to examine how accurate we can acheive in the estimated relation structure when a true model is known. The performance evaluation is discussed based on the use of FP, FN and the total error. In this section, we use $n = 10$ and generate A_{true} with random sparsity patterns. Then measurements Y are generated according to $Y = (I - A_{\text{true}})^{-1}e$ where e is normally distributed with variance of 0.1. We process the experiment as detail explained in Figure 5.1. Firstly, the sample covariance matrix (S) of measurement Y is computed and we choose a set of regularization parameters by $\gamma_i \in [\gamma_1, \ldots, \gamma_{\max}]$ where $\gamma_{\max}$ is the γ that penalizes all entries in A to become zero. For each value of γ, $\hat{A}$ is obtained from solving the sparse SEM and the estimated zero pattern of $\hat{A}$ is kept. In the next step, this zero pattern is used as the zero constraint, $P(A) = 0$ in the primal convex SEM. The estimate of A from this step is further evaluated by a model selection score. We repeat this process using all values of γ and obtain a set of candidate models. The model selection criterions, *e.g.*, BIC, AIC, AICc, KIC and KICc, are calculated on each model. The main results are explained as follows.

Firstly, there are four main aspects that could influence the estimation results. These factors are sparsity density of the true model (A_{true}), the number of sample sizes (N), the number of known

zero locations used in the estimation, and the choice of model selection scores. The experiments are then designed to investigate the effects of these factors which can be explained below.

1. The sparsity density of A_{true}. In this experiment, we generate A_{true} with two sparsity levels, 50% and 80% and observe a relation between the sparsity pattern of $\hat{A}$ that minimizes BIC score and the error rate. A typical result is illustrated in Figures 7.8 when A_{true} is dense and in Figures 7.10 when A_{true} is sparse. The result shows that when A_{true} is sparse, our exploratory SEM formulation provides less FP and FN than the case that A_{true} is dense. Unavoidable errors as FP and FN are commonly seen since these type of errors occur against the hypothesis of the true model. Moreover, when A_{true} is dense, Figure 7.11 shows that using AIC leads to the minimum total error since this score is prone to use a dense model (which agree with the assumption on the true model). Similarly, when A_{true} is sparse, Figure 7.12 confirms that using the scores penalizing more on the model complexity such as BIC, AICc, KICc yields a lower total error.

2. The number of samples. In the experiments, we use $N = 100$ (moderate size) and $N = 100,000$ (large sample size) to examine the asymptotic properties of the estimates. When N is large, Figures 7.7 and 7.9 confirm that the selected γ is closer to zero since the sparse SEM (as a regularized problem) should yield the solution closer to that of non-regularized problem. Moreover, Figures 7.8 and Figures 7.10 report that, with same sparsity level, FP does not significantly change, but FN obviously decreases when N increases. This effect is also shown in Figure 7.11 and Figure 7.12 that FP also increases, but FN decreases to zero, showing that our regularized formulation is robust to false negative errors.

3. The percentage of known zero locations in the estimation. To examine this factor, the experiments are performed with the percentage of known zeros of $0\%, 20\%$ and 50%. The first two values correspond to the problem with negative df where the regularized solution could be not unique implying that the estimated zero pattern may not be as accurate as when knowing more zero locations. Figures 7.8, 7.10, 7.11 and 7.12 show that when we know more about the true zero locations in A_{true}, FP decreases, but FN seems to be indifferent.

4. The choice of model selection scores. We considered AIC (tend to choose dense models), AICc (adjusted for finite sample size), BIC, KIC and KICc scores (tend to choose simpler models). From Figures 7.11 and 7.12, it is verified that AIC tends to yield the minimum total error when A_{true} is dense, and conversely, the choices of BIC, AICc and KICc tends to provide the total error lower than other criterions when A_{true} is sparse.

Secondly, we summarize the behaviours of the performance measures, *i.e.*, FP, FN and total error obtained from the experiments from Figures 7.11 and 7.12. These two plots are averaged from 50 runs of sample covariance matrix S. The discussions are explained as follows.

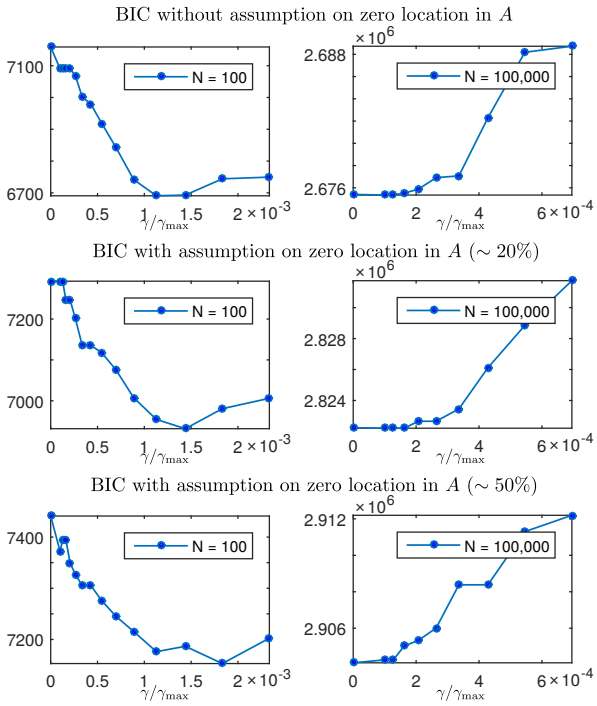


Figure 7.7: BIC scores as α varies when A_{true} **is dense**. This plot illustrates BIC scores when the sample of measurement (N) is 100 and $100,000$ according to three cases, *i.e.*, (top) no assumption of true zero location in A_{true}, (middle) knowing the true zero location in $A_{\text{true}} \sim 20\%$ and (bottom) knowing the true zero location in $A_{\text{true}} \sim 50\%$, in the estimation process.

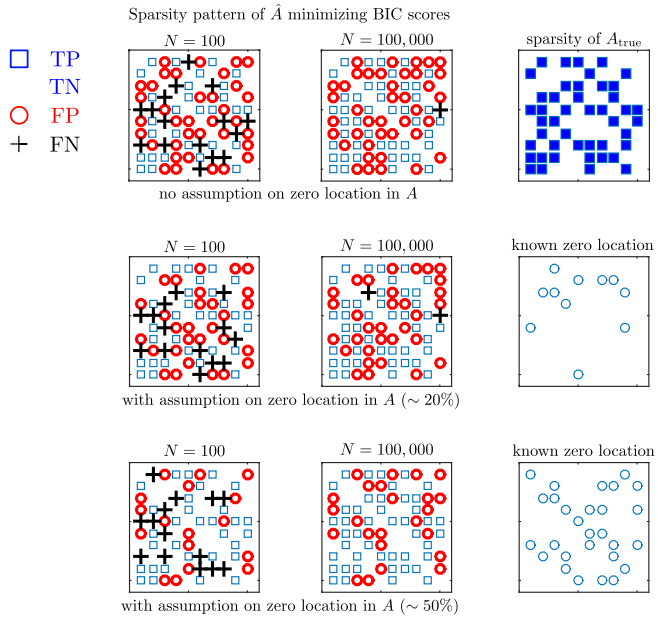


Figure 7.8: The sparsity pattern of $\hat{A}$ that minimizes BIC scores corresponding to Figure 7.7 when A_{true} **is dense**. This plot illustrates the sparsity pattern of $\hat{A}$ chosen via BIC scores when the sample of measurement (N) is 100 and $100,000$ according to three cases, *i.e.*, (top) no assumption of true zero location in A_{true}, (middle) knowing the true zero location in $A_{\text{true}} \sim 20\%$ and (bottom) knowing the true zero location in $A_{\text{true}} \sim 50\%$, in the estimation process.

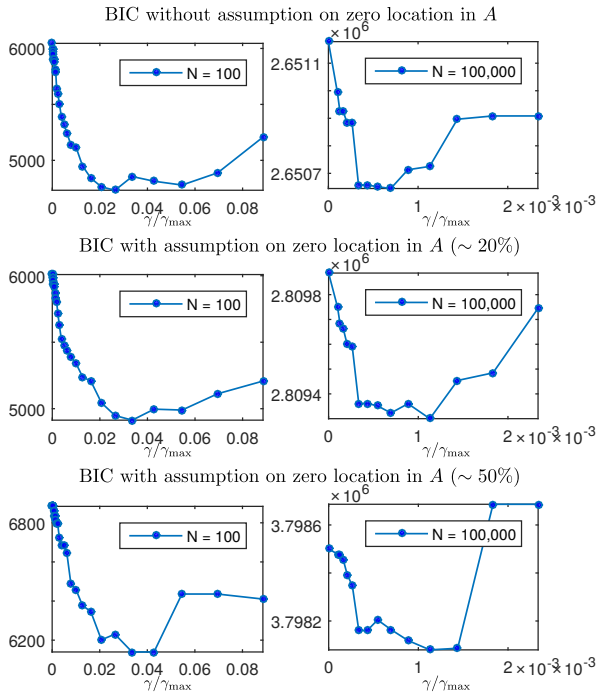


Figure 7.9: BIC scores as α varies when A_{true} **is sparse**. This plot illustrates BIC scores when the sample of measurement (N) is 100 and $100,000$ according to three cases, *i.e.*, (top) no assumption of true zero location in A_{true}, (middle) knowing the true zero location in $A_{\text{true}} \sim 20\%$ and (bottom) knowing the true zero location in $A_{\text{true}} \sim 50\%$, in the estimation process.

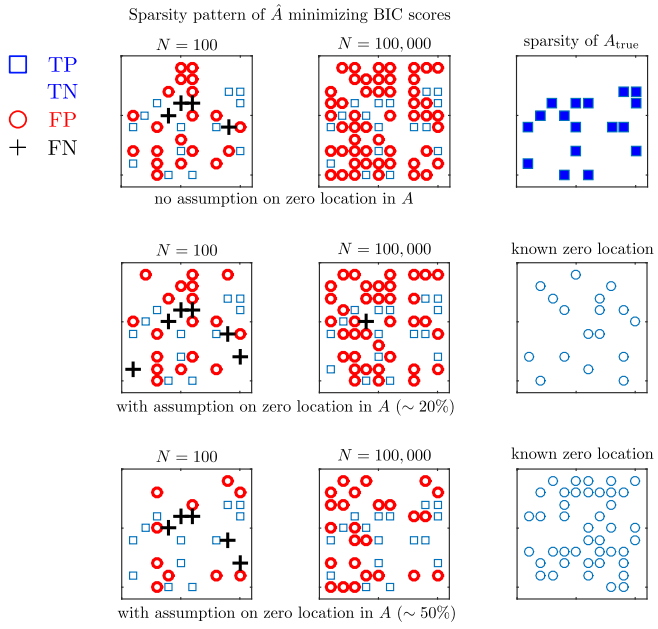
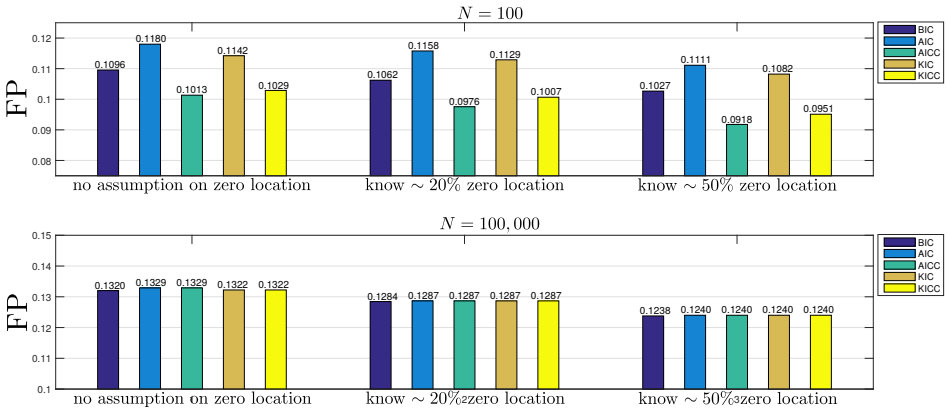
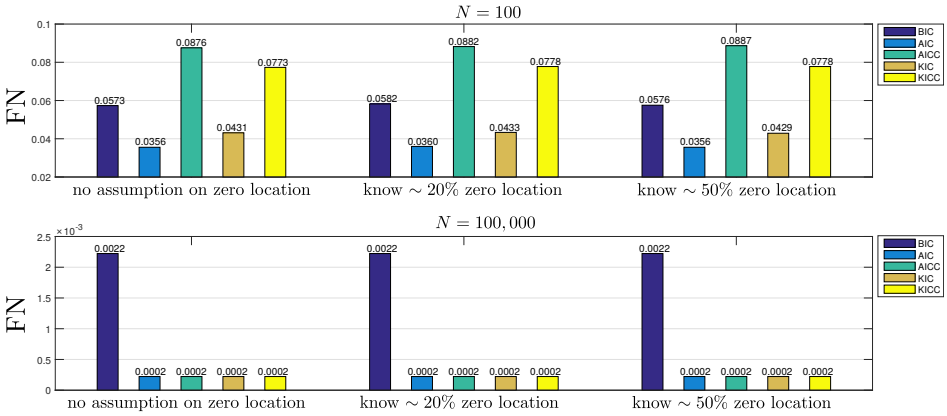


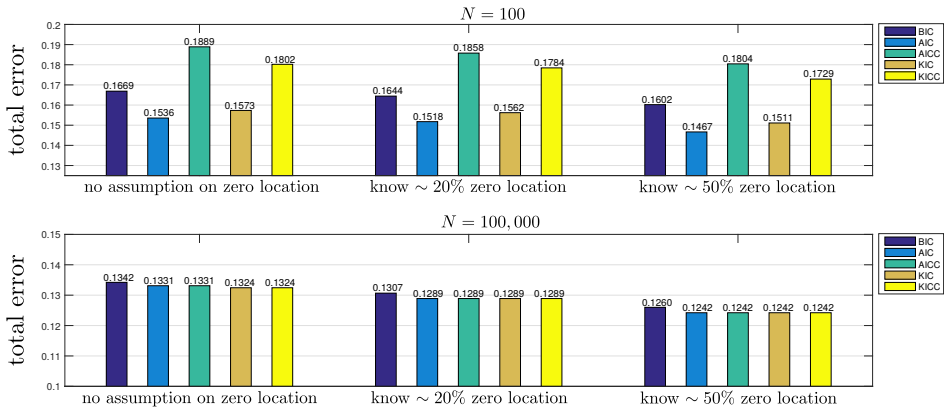
Figure 7.10: The sparsity pattern of $\hat{A}$ that minimizes BIC scores corresponding to Figure 7.9 when A_{true} **is sparse**. This plot illustrates the sparsity pattern of $\hat{A}$ chosen via BIC scores when the sample of measurement (N) is 100 and $100,000$ according to three cases, *i.e.*, (top) no assumption of true zero location in A_{true}, (middle) knowing the true zero location in $A_{\text{true}} \sim 20\%$ and (bottom) knowing the true zero location in $A_{\text{true}} \sim 50\%$, in the estimation process.



(a) **False Positive (FP) error.** When A_{true} is dense, FP from all model selection criterions tends to decrease when we use more knowledge about zero location in A_{true} into the estimation process, but it increases when N grows as all model selection criterions tend to select the denser $\hat{A}$. In the case of small N, AICc provides the minimum FP error.

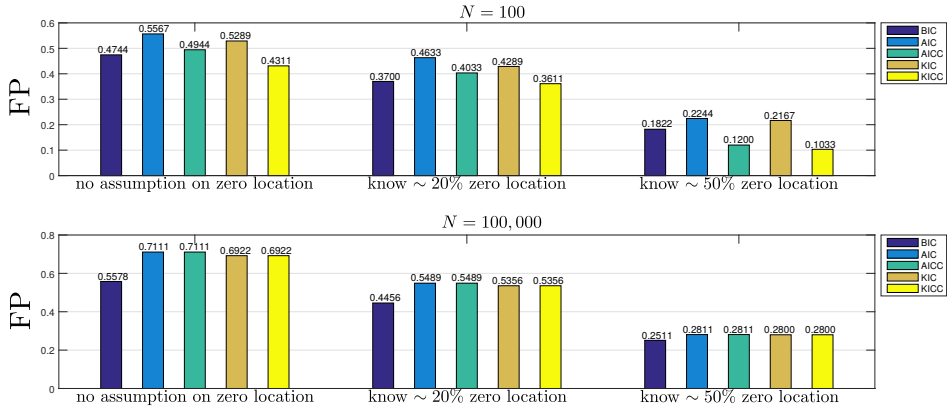


(b) **False Negative (FN) error.** Using more knowledge about zero location in A_{true} into the estimation process barely affects the change of FN, but it can be improved when N grows.

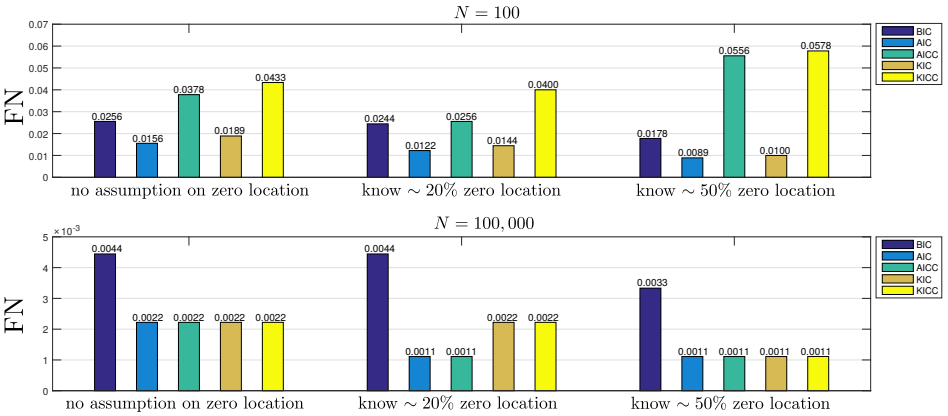


(c) **Total error.** Main portion of the total error comes from FP so it tends to decrease when we have more assumption about true zero location in A_{true}.

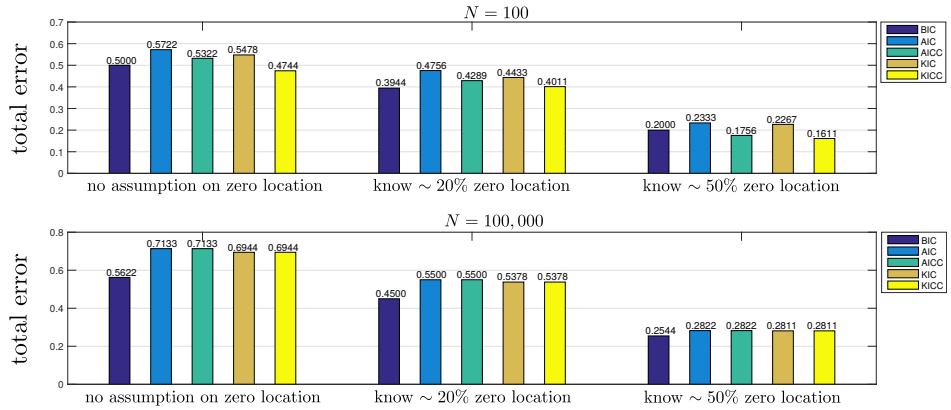
Figure 7.11: Averaged FP, FN and total error from 50 runs of sample covariance matrix S, when A_{true} **is dense.** The results show that AIC provides the mimum error when N is small.



(a) **False Positive (FP) error.** FP from all model selection criterions tends to decrease when we use the knowledge about zero location in A_{true} into the estimation process, but it increases when N grows as all model selection criterions tend to select a denser $\hat{A}$. In the case of small N, BIC, AICc and KICc provide the better accuracy as the true model is sparse.



(b) **False Negative (FN) error.** Using more knowledge about zero location in A_{true} into the estimation process barely affects the change of FN, but it can be improved when N grows.



(c) **Total error.** Main portion of the total error comes from FP so it tends to decrease when we have more assumption about true zero location in A_{true}.

Figure 7.12: Averaged FP, FN and total error from 50 runs of sample covariance matrix S, when A_{true} **is sparse.** The results show that BIC, AICc and KICc provide the lower total error when N is small.

1. False positive (FP). Regardless of the true sparsity level in A_{true}, FP from all model selection criterions tends to decrease when we know more true zero location in A_{true}. For large N, all model selection criterions tends to choose denser models so that FP highly increases. BIC, AICc and KICc provide better accuracies when N is small.

2. False negative (FN). Using more knowledge about true zero location in A_{true} hardly affects the improvement of FN regardless of the sparsity in A_{true}. But for large N, FN tends to decease to zero. In the case that we have small N, AIC and KIC provide better accuracy.

3. Total error. Total error highly depends on FP since the value of FN is always small comparing to FP so, this error from all model selection criterions tends to decrease when we use the assumption about true zero location in the estimation process. When N is small, in the case that A_{true} is dense, AIC provides the minimum total error. In contrast, BIC, AICc and KICc provide the better accuracies when A_{true} is sparse as they penalize more on the model complexity.

Finally, we would like to compare the entry magnitudes of $\hat{A}$ and A_{true}. We select one typical example in the experiment that A_{true} is sparse; the percentage of known zero locations is 50%; $N = 100,000$; and $\hat{A}$ is chosen by the BIC score. All entries in A_{true} are sorted from small to large and all entries in $\hat{A}$ are sorted following these indices. Figure 7.13 shows that when entries of A_{true} are nonzero, the estimated entries mostly have the same signs. When entries of A_{true} are zero, the errors in the estimated entries occur with mostly small magnitudes.

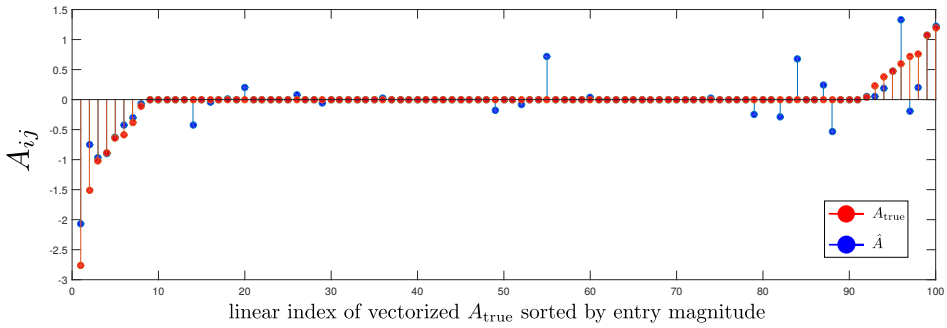


Figure 7.13: Similarity pattern of entry magnitudes between sparse A_{true} and $\hat{A}$ that minimizes BIC score. We use $N = 100,000$ and assumption to known $\sim 50\%$ zero locations in the estimation process. The result shows that the magnitude of each entry in $\hat{A}$ is quite equal to the magnitude of corresponding entry in A_{true}.

7.4 Algorithm performance

To see the algorithm performance for solving our both primal convex SEM and sparse SEM, we generate data with $n = 50, 100, \ldots, 1000$, using 50 samples of S for each n. We solve primal convex SEM (6.3) and sparse SEM (6.9) using ADMM as the details explained in chapter 6. Then we plot the averaged CPU time versus n. The computer's specification used in this experiment is:

CPU : Intel Core I5-6400 (2.7 GHz), RAM : 16GB DDR4 BUS2133, HDD : SATA III 7200 RPM (1TBs), OS : WINDOWS10-64bit Education. Solving either primal convex SEM or sparse SEM with dimension n involves total number of variables in X, $(n(n+1)/2$, plus the number of variables in Ψ and the number of of paths in A.

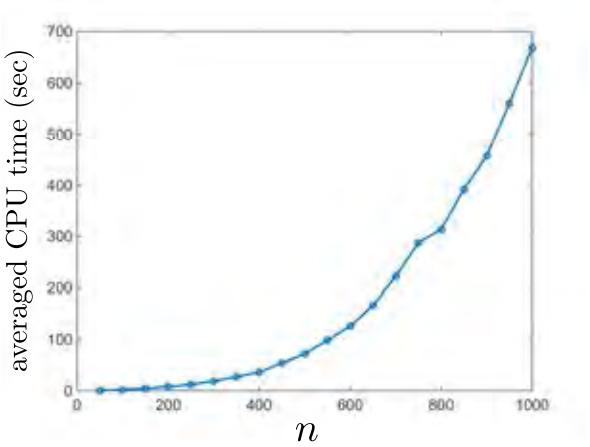


Figure 7.14: Averaged CPU time used to solve primal convex SEM from 50 samples of S for each n. With $n = 1000$, it takes around 11 minutes.

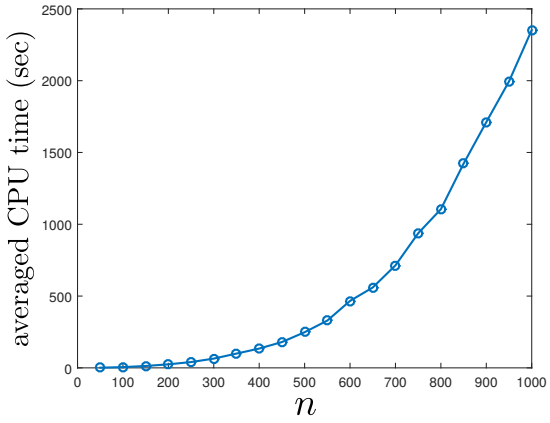


Figure 7.15: Averaged CPU time used to solve sparse SEM with ℓ_1-regularization from 50 samples of S for each n. With $n = 1000$, it takes around 40 minutes.

For solving both primal convex and sparse SEM, the main computational cost of ADMM algorithm only depends on eigenvalue decomposition of symmetric matrix with size $2n$, which is $\mathcal{O}((2n))^3$, where n is a number of variables. The averaged CPU time used is shown in Figure 7.14 and Figure 7.15. A trial problem with $n = 1000$ and a given pattern in A, resulting in totally $1,000,000$ variables, it requires approximately about 11 minutes for solving the primal convex SEM and about 42 minutes for solving the sparse SEM, respectively. A large-scale setting like this may not be feasible when implemented with an iterative method based on the use of Hessian matrix.

7.5 Learning causal relation among brain regions from exploratory SEM

In this experiment, we apply our framework to explore the causal relations among brain regions from fMRI data. This is an fMRI data set recorded with a fast sampling rate (sampling time 0.1 s) and from 21 subjects reported in [30]. The subjects were instructed to press a button responded to right or left visual hemi-field stimuli. For each time series of a subject, it contains signals from $1,100$ voxels with 299 time points where voxels are divided to belong to each of the 10 regions of interest (ROIs): left and right visual (V), left and right parietal (PCC), left and right premotor (PreM), left and right somatosensory (S) and left and right motor (M).

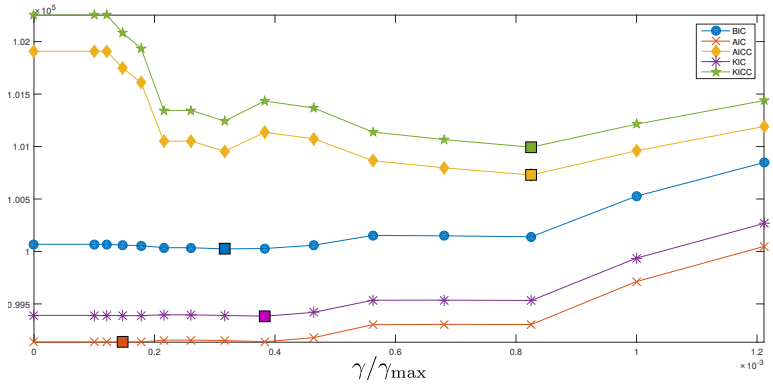


Figure 7.16: Scores of all model selection criterions: BIC, AIC, AICc, KIC and KICc. A square in each line corresponds to the minimum score of each criterion.

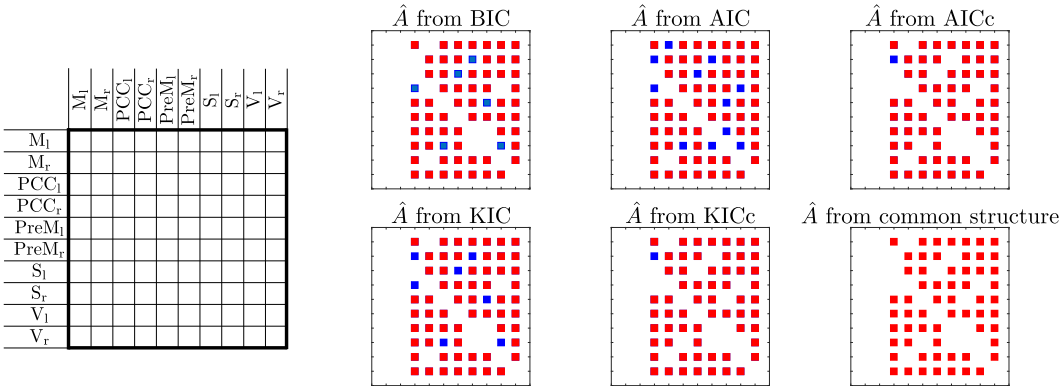


Figure 7.17: The sparsity pattern of $\hat{A}$ that is selected by each model selection criterion. The subscript *l* or *r* denote the corresponding ROIs that locate on the *left* or *right* hemisphere, respectively. The red squares represent a common pattern from all sparsity pattern in each $\hat{A}$. The first and second column of $\hat{A}$ are zero according to the assumption that motor area must be the end point of this brain network.

Firstly, the signals in each ROI are averaged in spatial domain resulting in 10 time series with 299 time points, corresponding to $Y \in \mathbf{R}^{10\times 299}$. The sample covariance matrix is computed from Y and a set of regularization parameters are selected as $\gamma_i \in \{\gamma_1, \ldots, \gamma_{\max}\}$ where $\gamma_{\max}$ is the γ that penalizes all entries in A to be zero. The experiment paradigm of this data [30] hypothetically suggested that the flows in the brain network should end in the motor area. In other words, there

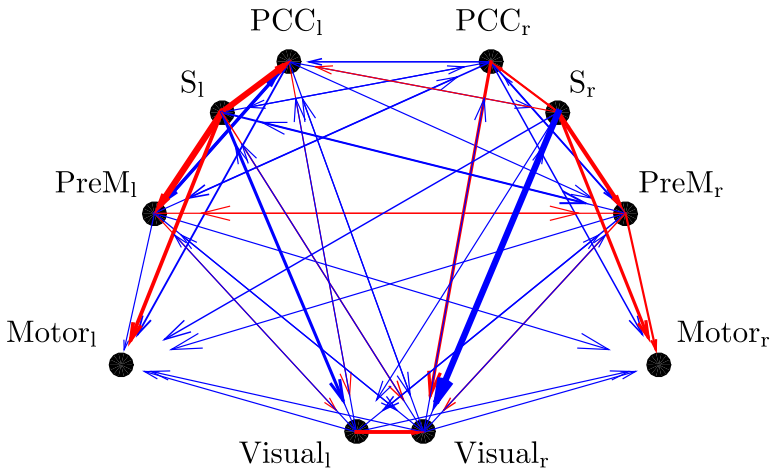


Figure 7.18: Brain structure from the common path matrix. This result shows the relation among visual (V), parietal (PCC), somatosensory (S), premotor (PreM) and motor (M) region of left and right hemisphere. The magnitude of path coefficients affects to line width. Positive and negative path coefficients are represented by red and blue color, respectively.

are no outgoing links from motor to the other areas. Therefore, we can model that the columns in A corresponding to the Motor area should be entirely zero and this can be encoded as the constraint $P(A) = 0$ in the sparse SEM problem. After solving sparse SEM with many γ, we obtain a set of candidate model, each of which is labeled by model selection scores: BIC, AIC, AICc, KIC and KICc. These scores are shown in Figure 7.16 as γ varies. Each line in this plot represents a score of one model selection criterion and the squares indicate the minimum score of each criterion. The common zero structure of $\hat{A}$'s (obtained by different model selection scores) shown in Figure 7.17 is used as the constraint $P(A) = 0$ and then the primal convex SEM subject to this constraint is solved, providing a low rank solution and so $\hat{\Sigma}$ is obtained by the inverse of the solution X. Finally, we obtain the structure of path matrix with optimal coefficients and this represents the structure of relation among brain ROIs as shown in Figure 7.18. The line widths in this plot are proportional to the magnitude of path coefficients and the colors which is red or blue represent the positive and negative of path coefficients, respectively.

Figure 7.18 shows an estimated brain network from one subject. It shows strong relations among somatosensory (S), parietal (PCC), premotor (PreM) and motor for both left and right hemisphere but for the relations in right hemisphere seem to be stronger. The strong connections also appear between somatosensory and visual. The ROIs on both sides of hemisphere seem to be free from each other as the relations between them are weak. Moreover, the pattern of connection in this network seems to be symmetric between left and right hemisphere. From the same fMRI data set, [30] has shown that improve the sampling rate can improve the sensitivity of Granger causality estimate. The directional causal influences between 5 ROIs in each hemisphere have been inferred that there are significant causal relations from visual to parietal, somatosensory, premotor and motor. For the left hemisphere, parietal has strong relations to somatosensory, motor and premotor but for the right

hemisphere, parietal only strongly affects to premotor. Similarly to [31], it has shown that the strong causal relation appears during the first 100 ms and these connections also start from visual to motor directly. The moderate relations also appear among visual, parietal, somatosensory and motor but the connections between ROIs of left and right hemisphere seem to be weak. From the result, the brain network from our findings is quite similar to those two works since the strong connection appears in the ROIs related to the motor task, and those ROIs also receive the strong influences from visual. However, this is a brain network concluded from one subject. An intensive experiment for brain network verification of several subjects can be considered in the future study. Finally, we comment that the casual relations of each brain ROI from our finding are considered to be only a contemporaneous effect.

CHAPTER VIII

CONCLUSION

This thesis has proposed the two estimation formulations for solving problems of path analysis in structural equation modeling (SEM). The first formulation, referred to as *the primal convex SEM*, can be an alternative method for solving a confirmatory SEM under conditions that i) the noise covariance in the model is homoskedastic, or that the covariance is a multiple of identity matrix, and ii) the estimated noise covariance is controlled by a problem parameter, α, chosen to be sufficiently small. We have shown by some proofs and extensive experiments that, for solving a primal convex SEM, the choice of $\alpha = \lambda_{\min}(S)$ is suggested to obtain a low rank solution which is useful as the estimated covariance of the model that can be chosen to be X^{-1} (the $(1,1)$ block of our solution matrix). The second formulation, denoted by *the sparse SEM*, is a regularized estimation proposed for exploratory SEM by adding ℓ_1-type penalty of the path coefficient matrix. This formulation is in the area of sparse estimation. We have derived the critical value of regularization parameters, $\gamma_{\max}$, that can enforce all entries in estimated path matrix to be zero. The expression of this value allow us to consider a model selection problem using a sufficient number of candidate models. We have shown the performance of our sparse SEM by experiments which show that its performance depends on i) the percentage of known zero locations of the true model in the estimation and ii) the number of sample sizes.

We also provide a scheme for learning causal relation structures among variables by applying both primal convex SEM and sparse SEM formulations which we refer to as *exploratory SEM*. The performance of our exploratory SEM has been evaluated from the simulation process showing that there are four factors, *i.e.*, the density level of the true model, the percentage of known zero locations of the true model in the estimation, the number of sample sizes and the choice of model selection scores, can affect to this performance. Another important result from this experiment illustrates that if the causal structure of true model is complex, AIC provides the minimum total error. In contrast, if the causal structure of true model is simpler, BIC, AICc and KICc provide the better accuracy. An application of this scheme was preliminarily illustrated by learning causal relations among brain regions from fMRI data. The brain network illustrating the causal relations among ROIs says that there are strong connections among somatosensory, parietal, premotor and motor for both left and right hemisphere. In particular, the dominant pairs of strong connection are somatosensory$\rightarrow$visual, somatosensory$\rightarrow$parietal and somatosensory$\rightarrow$premotor. The brain network we found practically agreed with the brain network from the original paper used this data set. Finally, we comment that the relation structures from our scheme are only considered to be a contemporaneous effect.

Another contribution of this thesis is the numerical method based on ADMM algorithm that is suitable for solving the two formulations in a large-scale setting. The main computational cost in each step highly depends on the cost of eigenvalue decomposition of a symmetric matrix with size $2n$ where n is the problem dimension.

Throughout this thesis, we state that the solution from our estimation formulation is meaningful when it has low rank and the occurrence of this solution has been shown from experiments. Therefore an area of our future study is to provide the analytical result of conditions on problem parameters that can guarantee to obtain this low rank solution. We observed that when α which is our input parameter is small enough, the low rank solution holds and when this holds, rank of the dual variable, $Z \in \mathbf{S}^{2n}$, is n at optimum. If we consider the dual problem under an assumption that Z has low rank, its optimality condition turns out to be the nonlinear matrix equation. Hence we may find the condition depending on the range of α that is necessary to have the low rank solution by examining the existence of positive definite solution of the dual problem.

Bibliography

[1] R. MacCallum and J. Austin, "Applications of structural equation modeling in psychological research," *Annual review of psychology*, vol. 51, no. 1, pp. 201–226, 2000.

[2] H. Hassan and M. Abdel-Aty, "Analysis of drivers' behavior under reduced visibility conditions using a structural equation modeling approach," *Transportation research part F*, vol. 14, pp. 614–625, 2011.

[3] N. Suki, "Passenger satisfaction with airline service quality in Malaysia: a structural equation modeling approach," *Research in transportation business & management*, vol. 10, pp. 26–32, 2014.

[4] U. Ramanathan and L. Muyldermans, "Identifying the underlying structure of demand during promotions: a structural equation modeling approach," *Expert systems with applications*, vol. 38, pp. 5544–5552, 2011.

[5] K. Bollen, *Structural equations with latent variables*. John Wiley & Sons, 1989.

[6] A. Mclntosh and F. Gonzalez-Lima, "Structural equation modeling and its application to network analysis in functional brain imaging," *Human Brain Mapping*, vol. 2, no. 1-2, pp. 2–22, 1994.

[7] C. Büchel and K. Friston, "Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI.," *Cerebral cortex*, vol. 7, no. 8, pp. 768–778, 1997.

[8] E. Bullmore, B. Horwitz, G. Honey, M. Brammer, S. Williams, and T. Sharma, "How good is good enough in path analysis of fMRI data?," *NeuroImage*, vol. 11, no. 4, pp. 289–301, 2000.

[9] J. Kim, W. Zhu, L. Chang, P. Bentler, and T. Ernst, "Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data," *Human Brain Mapping*, vol. 28, no. 2, pp. 85–93, 2007.

[10] G. Chen, D. Glen, Z. Saad, J. Hamilton, M. Thomason, I. Gotlib, and R. Cox, "Vector autoregression, structural equation modeling, and their synthesis in neuroimaging data analysis," *Computers in biology and medicine*, vol. 41, pp. 1142–1155, 2011.

[11] T. Raykov and G. Marcoulides, *A First Course in Structural Equation Modeling*. Lawrence Eribaum Associates, Inc., second ed., 2006.

[12] K. Jöreskog, D. Sörbom, S. D. Toit, and M. D. Toit, *LISREL 8: New statistical features*. Scientific Software International, 2000.

[13] S. Mulaik, *Linear causal modeling with structural equations*. CRC Press, 2009.

[14] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer, 2nd ed., 2009.

[15] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.

[16] R. Hoyle, *Structural equation modeling: Concepts, issues, and applications*. Sage Publications, 1995.

[17] M. Grant and S. Boyd, *CVX: Matlab software for disciplined convex programming (web page and software)*. `http://stanford.edu/~boyd/cvx`, 2007.

[18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004. Available: `www.stanford.edu/~boyd/cvxbook`.

[19] R. Jacobucci, K. Grimm, and J. McArdle, "Regularized structural equation modeling," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 00, pp. 1–12, 2016.

[20] L. Vandenberghe, "Convex optimization techniques in system identification," in *Proceedings of the IFAC Symposium on SystemIdentification*, pp. 71–76, 2012.

[21] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2010.

[22] F. Lin, M. Fardad, and M. Jovanovic, "Design of optimal sparse feedback gains via the alternating direction method of multipliers," *IEEE Transactions on Automatic Control*, vol. 58, no. 9, pp. 2426–2431, 2013.

[23] H. Akaike, "Factor analysis and AIC," *Psychometrika*, vol. 52, no. 3, pp. 317–332, 1987.

[24] H. Akaike, "Akaike's Information Criterion," in *International Encyclopedia of Statistical Science*, pp. 25–25, Springer, 2011.

[25] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*, vol. 1. Springer series in statistics Springer, Berlin, 2001.

[26] C. Hurvich and C. Tsai, "A Corrected Akaike Information Criterion for vector autoregressive model selection," *Journal of time series analysis*, vol. 14, no. 3, pp. 271–279, 1993.

[27] J. Cavanaugh, "A Large-sample model selection criterion based on Kullback's symmetric divergence," *Statistics & Probability Letters*, vol. 42, no. 4, pp. 333–343, 1999.

[28] A. Seghouane, "Vector autoregressive model-order selection from finite samples using Kullback's symmetric divergence," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 53, no. 10, pp. 2327–2335, 2006.

[29] E. Alpaydin, *Introduction to machine learning*. MIT press, 2014.

[30] F. Lin, J. Ahveninen, T. Raij, T. Witzel, Y. Chu, I. Jääskeläinen, K. Tsai, W. Kuo, and J. Belliveau, "Increasing fMRI sampling rate improves granger causality estimates," *Psychological Methods*, vol. 17, no. 1, pp. 1–14, 2012.

[31] E. Karahan, P. Rojas-Lopez, M. Bringas-Vega, P. Valdes-Hernandez, and P. Valdes-Sosa, "Tensor analysis and fusion of multimodal brain images," *Proceedings of the IEEE*, vol. 103, no. 9, pp. 1531–1559, 2015.

[32] R. Rockafellar and R.-B. Wets, *Variational Analysis*, vol. 317. Springer Science & Business Media, 2009.

[33] N. Parikh and S. Boyd, "Proximal algorithms," *Foundations and Trends in Optimization*, vol. 1, no. 3, pp. 127–239, 2014.

APPENDIX

APPENDIX

In this chapter, we provide technical derivations of the results used in the thesis.

9.1 Dual problem of the primal convex SEM

In this section we show that the dual of (3.1) which is the problem

$$\begin{array}{ll} \text{minimize} & -\log\det X + \mathbf{tr}(SX), \\ \text{subject to} & \begin{bmatrix} X & (I-A)^T \\ I-A & \Psi \end{bmatrix} \succeq 0, \\ & 0 \preceq \Psi \preceq \alpha I, \\ & P(A) = 0, \end{array} \tag{9.1}$$

with variables $X \in \mathbf{S}^n$, $A \in \mathbf{R}^{n\times n}$ and $\Psi \in \mathbf{S}^n$, is the one given in (3.3)

$$\begin{array}{ll} \text{minimize} & -\log\det(S - Z_1) - 2\,\mathbf{tr}(Z_2) - \alpha\,\mathbf{tr}(Z_4) + n, \\ \text{subject to} & Z = \begin{bmatrix} Z_1 & Z_2^T \\ Z_2 & Z_4 \end{bmatrix} \succeq 0, \\ & Q(Z) = 0, \end{array} \tag{9.2}$$

with variable $Z \in \mathbf{S}^{2n}$.

Derivation of the dual problem. Let $Z = \begin{bmatrix} Z_1 & Z_2^T \\ Z_2 & Z_4 \end{bmatrix} \in \mathbf{S}^{2n}$, $\Omega \in \mathbf{S}^n$ and $U \in \mathbf{R}^{n\times n}$ be the Lagrange multipliers of the constraints

$$\begin{bmatrix} X & (I-A)^T \\ I-A & \Psi \end{bmatrix} \succeq 0, \quad \Psi \preceq \alpha I, \quad P(A) = 0,$$

respectively. The Lagrangian of the problem (3.1) is

$$\begin{aligned} L(X, A, \Psi, Z, \Omega, U) = &-\log\det X + \mathbf{tr}(SX) - \mathbf{tr}(Z_1 X) - 2\,\mathbf{tr}(Z_2) + 2\,\mathbf{tr}(Z_2^T A) \\ &- \mathbf{tr}(Z_4\Psi) + \mathbf{tr}(\Omega\Psi) - \alpha\,\mathbf{tr}(\Omega) - 2\,\mathbf{tr}(U^T P(A)). \end{aligned} \tag{9.3}$$

The infimum of L with respect to the primal variables can be determined as follows.

- The term in L that is a function of Ψ is $\mathbf{tr}(\Omega\Psi) - \mathbf{tr}(Z_4\Psi)$. This function is linear in Ψ, so the infimum of L with respect to Ψ exists (and is zero) if

$$\Omega = Z_4. \tag{9.4}$$

- The term in L that is a function of X is given by $-\log\det X + \mathbf{tr}(SX) - \mathbf{tr}(Z_1 X)$ which can be minimized when its gradient with respect to X is zero. This gives

$$-X^{-1} + S - Z_1 = 0 \tag{9.5}$$

  or that $X = (S - Z_1)^{-1}$ and

$$\inf_X \ \{-\log\det X + \mathbf{tr}(SX) - \mathbf{tr}(Z_1 X)\} = \log\det(S - Z_1) + n.$$

- Lastly, the infinum of the term in L that is a function of A (up to the scaling factor 2) is given by

$$\inf_A \left\{\mathbf{tr}(Z_2^T A) - \mathbf{tr}(U^T P(A)\right\} \quad = \quad \inf_A \left\{\mathbf{tr}(Z_2^T A) - \mathbf{tr}(P(U)^T A)\right\}$$

  where we have used the fact that the operator P defined in (2.3) is self-adjoint, *i.e.*, $\mathbf{tr}(U^T P(A)) = \mathbf{tr}(P(U)^T A)$. Hence, the expression is linear in A, so the infimum is zero provided that $Z_2 = P(U)$. This means the (i,j) entries of Z_2 for $(i,j) \in I_A$ are free variables, and the other entries of Z_2 must be zero. This can be written in the matrix format as $P^c(Z_2) = 0$.

The minimized Lagrangian with respect to the primal variables provides us the dual function

$$g(Z) = \log\det(S - Z_1) + n - 2\,\mathbf{tr}(Z_2) - \alpha\,\mathbf{tr}(Z_4)$$

with the domain constraints:

$$Z \succeq 0, \quad P^c(Z_2) = 0.$$

The last constraint on Z is equivalent to $Q(Z) = 0$ (recall the definition of Q in (2.5)). The dual is the problem of maximizing the dual function which is obtained directly.

9.2 Subgradients and subgradient calculus

We have seen that the convex formulation for the sparse SEM (4.1) is a nondifferentiable problem. Its cost objective is nondifferentiable and hence the gradient does not exist at $A = 0$. In this section, we present the generalized concept of gradient for nondifferentiable functions called *subgradients* [32].

Let $f : \mathbf{R}^n \to (-\infty, \infty)$ be a convex function and let $z \in \mathbf{dom}\, f$. An element of $g \in \mathbf{R}^n$ is called a subgradient of f at z if

$$\langle g, x - z\rangle \leq f(x) - f(z)$$

for all $x \in \mathbf{R}^n$. A subgradient of f at z might not be unique, so the set of all possible subgradients of f at z is called the *subdifferential* of f at z denoted by $\partial f(z)$.

Let us provide an example of subgradients of the $1 - norm$ function which will be used often in this thesis. Consider $f(x) = \|x\|_1, x \in \mathbf{R}^n$. The subgradient of f and $x = 0$ is given by

$$\partial f(0) = g \in \mathbf{R}^n, \quad \text{where} \quad \|g\|_\infty \leq 1.$$

To verify this result, we check from the definition of the subgradient that g must satisfy $\langle g, x - 0\rangle \leq f(x) - f(0)$ for all $x \in \mathbf{R}^n$ which can be rewritten as

$$g^T x \leq \|x\|_1, \quad \forall x \in \mathbf{R}^n. \tag{9.6}$$

If (9.6) holds for all x then setting $x = \pm e_i$ (the standard unit vector) gives $|g_i| \leq 1$. Therefore, a subgradient of $\|x\|_1$ at $x = 0$ can be any vector g that $\|g\|_\infty \leq 1$. We can conclude easily from a special of this example that a subgradient of $f(x) = |x|$ at $x = 0$ is $\partial f(0) = g$ where $g \in \mathbf{R}$ is any scalar that $|g| \leq 1$.

Optimality condition for nonsmooth problems. Consider an optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(x), \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m. \end{array}$$

where f_i is convex, defined on $\mathbf{R}^n$ and subdifferentiable for $i = 0, 1, \ldots, m$. Moreover the strict feasibility (or Slater's condition) is assumed. Then, $x^\star$ is primal optimal ($\lambda^\star$ is dual optimal) if and only if

- primal feasibility: $f_i(x^\star) \leq 0$ for $i = 1, \ldots, m$
- dual feasibility: $\lambda^\star \geq 0$
- zero is in the subdifferential of the Lagrangian

$$0 \in \partial f_0(x^\star) + \sum_{i=1}^m \lambda_i^\star \partial f_i(x^\star)$$

- complementary slackness condition: $\lambda_i^\star f_i(x^\star) = 0$ for $i = 1, 2, \ldots, m$

The above conditions are known as KKT conditions for *nondifferentiable* f_i for $i = 0, 1, \ldots, m$.

9.3 Dual problem of the sparse SEM

In Section 4, we have stated that the dual of the sparse SEM problem (4.1) which is given by

$$\begin{array}{ll} \text{minimize} & -\log\det X + \mathbf{tr}(SX) + 2\gamma \sum\limits_{(i,j)\notin I_A} |A_{ij}|, \\ \text{subject to} & \begin{bmatrix} X & (I-A)^T \\ I-A & \Psi \end{bmatrix} \succeq 0, \\ & 0 \preceq \Psi \preceq \alpha I, \\ & P(A) = 0, \end{array} \tag{9.7}$$

with variables $X \in \mathbf{S}^n$, $A \in \mathbf{R}^{n\times n}$ and $\Psi \in \mathbf{S}^n$, is the problem

$$\begin{array}{ll} \text{maximize} & \log\det(S - Z_1) - 2\,\mathbf{tr}(Z_2) - \alpha\,\mathbf{tr}(Z_4) + n, \\ \text{subject to} & \begin{bmatrix} Z_1 & Z_2^T \\ Z_2 & Z_4 \end{bmatrix} \succeq 0, \\ & |(Z_2)_{ij}| \le \gamma, \quad \forall\,(i,j) \notin I_A, \end{array} \tag{9.8}$$

with variable $Z \in \mathbf{S}^{2n}$.

In this Appendix, we provide the details of the dual problem derivation and the KKT conditions.

Derivation of the dual problem. Let $Z = \begin{bmatrix} Z_1 & Z_2^T \\ Z_2 & Z_4 \end{bmatrix} \in \mathbf{S}^{2n}$, $\Omega \in \mathbf{S}^n$ and $U \in \mathbf{R}^{n\times n}$ be the Lagrange multipliers of the constraints

$$\begin{bmatrix} X & (I-A)^T \\ I-A & \Psi \end{bmatrix} \succeq 0, \quad \Psi \preceq \alpha I, \quad P(A) = 0$$

respectively. With the notation

$$h(A) = \sum_{(i,j)\notin I_A} |A_{ij}|, \tag{9.9}$$

the Lagrangian of (9.7) is

$$\begin{aligned} L(X, A, \Psi, Z, \Omega, U) = &- \log\det X + \mathbf{tr}(SX) + 2\gamma h(A) - \mathbf{tr}(Z_1 X) - 2\,\mathbf{tr}(Z_2) + 2\,\mathbf{tr}(Z_2^T A) \\ &- \mathbf{tr}(Z_4\Psi) + \mathbf{tr}(\Omega\Psi) - \alpha\,\mathbf{tr}(\Omega) - 2\,\mathbf{tr}(U^T P(A)). \end{aligned} \tag{9.10}$$

The infimum of L with respect to the primal variables can be determined as follows.

- The term in L that is a function of Ψ is $\mathbf{tr}(\Omega\Psi) - \mathbf{tr}(Z_4\Psi)$. This function is linear in Ψ, so the infimum of L with respect to Ψ exists (and is zero) if

$$\Omega = Z_4. \tag{9.11}$$

- The term in L that is a function of X is given by $-\log\det X + \mathbf{tr}(SX) - \mathbf{tr}(Z_1 X)$ which can be minimized when its gradient with respect to X is zero. This gives

$$-X^{-1} + S - Z_1 = 0 \tag{9.12}$$

  or that $X = (S - Z_1)^{-1}$ and

$$\inf_X \; \{-\log\det X + \mathbf{tr}(SX) - \mathbf{tr}(Z_1 X)\} = \log\det(S - Z_1) + n.$$

- Lastly, by using the concept of conjugate function in Appendix 9.5, the infinum of the term in

L that is a function of A is given by

$$\begin{aligned}\inf_A \left\{2\gamma h(A) + 2\,\mathbf{tr}(Z_2^T A) - 2\,\mathbf{tr}(U^T P(A)\right\} &= \inf_A \left\{2\gamma h(A) + 2\,\mathbf{tr}(Z_2^T A) - 2\,\mathbf{tr}(P(U)^T A)\right\} \\ &= -2\gamma \sup_A \left\{-h(A) - \mathbf{tr}\left(\left(\frac{Z_2 - P(U)}{\gamma}\right)^T A\right)\right\} \\ &= -2\gamma h^*\left(-\frac{Z_2 - P(U)}{\gamma}\right) \\ &= 0\end{aligned}$$

provide that

$$P(Z_2 - P(U)) = 0, \quad \|P^c(Z_2 - P(U))\|_\infty \leq \gamma$$

(see Proposition 3 and (9.27)). Since the entries in U can be chosen arbitrarily, and we have $P(Z_2 - P(U)) = P(Z_2) - P(U) = 0$, meaning that $P(Z_2)$ contains free entries. Moreover, $P^c(Z_2 - P(U)) = P^c(Z_2) - P^c(P(U)) = P^c(Z_2)$. We then conclude that the infimum of L with respect to A is zero when

$$\|P^c(Z_2)\|_\infty \leq \gamma, \tag{9.13}$$

or equivalently that $|(Z_2)_{ij}| \leq \gamma$ for $(i,j) \notin I_A$.

The minimized Lagrangian with respect to the primal variables provides us the dual function

$$g(Z) = \log\det(S - Z_1) + n - 2\,\mathbf{tr}(Z_2) - \alpha\,\mathbf{tr}(Z_4)$$

with the domain constraints:

$$Z \succeq 0, \quad |(Z_2)_{ij}| \leq \gamma, \quad \text{for all } (i,j) \notin I_A.$$

The dual problem is the problem of maximizing the dual function which is obtained directly.

Additionally, the conditions (9.12) and (9.11) are the condition of zero gradient of the Lagrangian with respect to X and Ψ (since L is differentiable with respect to these two). Moreover, from subgradient calculus, the condition (9.13) can be viewed as the condition that zero must be the one of subgradients of the Lagrangian, since L is nondifferentiable with respect to A. To show the latter, consider the term in L that involves only A (up to the scaling factor 2) which is equal to

$$\gamma \sum_{(i,j)\notin I_A} |A_{ij}| + \mathbf{tr}(Z_2^T A) - \mathbf{tr}(U^T P(A)) = \gamma \sum_{(i,j)\notin I_A} |A_{ij}| + \mathbf{tr}(Z_2^T A) - \mathbf{tr}(P(U)^T A)$$

by the self-adjoint property of the projection operator P. This term can be further rearranged as

$$\gamma \sum_{(i,j)\notin I_A} (|A_{ij}| + (Z_2)_{ij} A_{ij}) + \sum_{(i,j)\in I_A} ((Z_2)_{ij} A_{ij} - P(U)_{ij} A_{ij}) \tag{9.14}$$

by splitting the sum into two terms, and we can find subgradients of the above expression with respect A_{ij} separately. The gradient with respect to A_{ij} for $(i,j) \in I_A$ is $(Z_2)_{ij} - P(U)_{ij}$. Hence, zero

gradient condition gives $(Z_2)_{ij} = P(U)_{ij}$ for $(i,j) \in I_A$. This means the entries in Z_2 are free for $(i,j) \in I_A$. The subgradient of (9.14) with respect to A_{ij} for $(i,j) \notin I_A$ is given by

$$\gamma g_{ij} + (Z_2)_{ij}$$

where g_{ij} denotes a subgradient of the function $f(x) = |x|$ and g_{ij} is any number that $|g_{ij}| \leq 1$. The optimality condition that zero must be one of the subgradients gives $0 = \gamma g_{ij} + (Z_2)_{ij}$ for $(i,j) \notin I_A$. Since $|g_{ij}| \leq 1$, we have

$$\gamma \geq |(Z_2)_{ij}|, \quad \forall (i,j) \notin I_A$$

which is equivalent to the matrix notation $\gamma \geq \|P^c(Z_2)\|_\infty$ in (9.13). In conclusion, the optimality conditions can be presented in the following.

KKT conditions for exploratory SEM. If strong duality holds, then X, A, Ψ and Z are optimal if and only if the following conditions hold.

- Primal feasibility:

$$\begin{aligned} (I-A)^T\Psi^{-1}(I-A) &\preceq X, && (9.15)\\ 0 \prec \Psi &\preceq \alpha I, && (9.16)\\ P(A) &= 0. && (9.17) \end{aligned}$$

- Dual feasibility:

$$\begin{aligned} Z = \begin{bmatrix} Z_1 & Z_2^T \\ Z_2 & Z_4 \end{bmatrix} &\succeq 0, && (9.18)\\ \gamma &\geq \|P^c(Z_2)\|_\infty. && (9.19) \end{aligned}$$

- Zero gradient of the Lagrangian:

$$X = (S - Z_1)^{-1}. \quad (9.20)$$

- Complementary slackness condition:

$$\begin{aligned} \mathbf{tr}\left(\begin{bmatrix} Z_1 & Z_2^T \\ Z_2 & Z_4 \end{bmatrix}\begin{bmatrix} X & (I-A)^T \\ I-A & \Psi \end{bmatrix}\right) &= 0, && (9.21)\\ \mathbf{tr}\left(Z_4(\Psi - \alpha I)\right) &= 0. && (9.22) \end{aligned}$$

9.4 Derivation of $\gamma_{\max}$

The section 4 presents a convex formulation for estimating SEM model in (4.1) where the sparsity of the optimal path coefficient A can be controlled via the value of the penalty parameter, γ. In this section, we will show that there exists a critical value of γ, denoted by $\gamma_{\max}$ such that if

$$\gamma \geq \gamma_{\max}$$

then the optimal solution of A in 4.1 is the zero matrix. This means it is unreasonably recommended to increase γ arbitrarily in the problem, and we can use $\gamma_{\max}$ as an upperbound of the range of γ used in order to vary the sparsity patterns of A.

The derivation of $\gamma_{\max}$ is in fact derived from one of the optimality conditions of (4.1) in (9.19) and it is derived under an *assumption* that the optimal primal solution is *low rank*.

If the optimal primal solution has low rank, *i.e.*,

$$\mathbf{rank}\left(\begin{bmatrix} X & (I-A)^T \\ I-A & \Psi \end{bmatrix}\right) = n$$

then it follows from (9.21) that $\mathbf{rank}(Z) = n$ and $\mathbf{rank}(Z_4) = n$, so Z_4 is invertible. This further implies from (9.22) that $\Psi = \alpha I$. Since we aim to characterize the condition (9.19) when we obtain the sparsest solution of A, we set $A = 0$ in the optimal condition, then the matrix

$$\begin{bmatrix} X & (I-A)^T \\ I-A & \Psi \end{bmatrix} = \begin{bmatrix} X & I \\ I & \alpha I \end{bmatrix}$$

has rank n if and only if $X = (1/\alpha)I$. From (9.20), $Z_1 = S - X^{-1} = S - \alpha I$. Substitute this in the slackness condition (9.21)

$$\begin{bmatrix} X & I \\ I & \alpha I \end{bmatrix} \begin{bmatrix} S - \alpha I & Z_2^T \\ Z_2 & Z_4 \end{bmatrix} = 0,$$

we can solve for Z_2 as

$$Z_2 = \frac{1}{\alpha}(\alpha I - S),$$

and (9.13) becomes

$$\gamma \geq \frac{1}{\alpha}\|P^c(\alpha I - S)\|_\infty. \tag{9.23}$$

In conclusion, we have shown that if $A = 0$ is the optimal solution to (4.1) and the optimal primal solution has rank n, then (9.23) must be fulfilled. As a result, we can set

$$\gamma_{\max} = \frac{1}{\alpha}\|P^c(\alpha I - S)\|_\infty \tag{9.24}$$

as the critical value of γ, in the sense that for any $\gamma \geq \gamma_{\max}$, the optimal solution A must be zero.

9.5 Conjugate functions

In this part, we provide the derivation of the conjugate function for the functions considered in this thesis. We recall the definition of the conjugate function as follows.

$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x) \right).$$

Proposition 3. *Let $X \in \mathbf{R}^{n\times n}$ and $I_A \subseteq \{1,\ldots,n\}\times\{1,\ldots,n\}$ be an index set. Let $f : \mathbf{R}^{n\times n} \to \mathbf{R}$ be defined by $f(X) = \sum_{(i,j)\notin I_A} |X_{ij}|$. The conjugate function of f is*

$$f^*(Y) = \begin{cases} 0, & Y_{ij} = 0 \textit{ for } (i,j) \in I_A \textit{ and } \max\limits_{(i,j)\notin I_A} |Y_{ij}| \leq 1 \\ \infty, & \textit{otherwise}. \end{cases}$$

Proof. From the definition of the conjugate function

$$f^*(Y) = \sup_X \ (\mathbf{tr}(Y^T X) - f(X)), \tag{9.25}$$

we then characterize the term

$$\mathbf{tr}(Y^T X - f(X)) = \sum_{ij} Y_{ij}X_{ij} - \sum_{(i,j)\notin I_A} |X_{ij}| = \sum_{(i,j)\notin I_A} (Y_{ij}X_{ij} - |X_{ij}|) + \sum_{(i,j)\in I_A} Y_{ij}X_{ij}. \tag{9.26}$$

We first note that if $Y_{ij} \neq 0$ for $(i,j) \in I_A$ then the supremum does not exist because we can set $X_{ij} = 0$ for $(i,j) \notin I_A$ and choose $X_{ij} = t\mathbf{sign}(Y_{ij})$ for $(i,j) \in I_A$ and let $t \to \infty$. Therefore, the condition $Y_{ij} = 0$ for $(i,j) \in I_A$ must be in the domain of $f^*(Y)$. Setting this condition in (9.26) gives

$$\begin{aligned} \sup_X \ (\mathbf{tr}(Y^T X) - f(X)) &\leq \sup_X \sum_{(i,j)\notin I_A} |Y_{ij}||X_{ij}| - |X_{ij}| \\ &\leq \sup_X \sum_{(i,j)\notin I_A} \left(\max_{ij} |Y_{ij}| - 1 \right) |X_{ij}|. \end{aligned}$$

If $\max_{(i,j)\notin I_A} |Y_{ij}| > 1$ then the supremum does not exist again because we can choose X to be zero in all entries except that $X_{ij} = tY_{ij}$ for (i,j)th entry that corresponds to the maximum $|Y_{ij}|$ and let $t \to \infty$. If $\max_{(i,j)\notin I_A} |Y_{ij}| = 1$ then we have that

$$\sup_X \ (\mathbf{tr}(Y^T X) - f(X)) \leq 0,$$

i.e., $f^*(Y)$ is bounded above by zero. Therefore, if $\max_{(i,j)\notin I_A} |Y_{ij}| < 1$, we can achieve the upper bound by choosing $X = 0$ and $f^*(Y) = 0$ only when $Y_{ij} = 0, \forall (i,j) \in I_A$ and $\max_{(i,j)\notin I_A} |Y_{ij}| \leq 1$.

If we apply the notation of projection P in (2.3) and the definitions of matrix norms: $\|X\|_1 = \sum_{ij} |X_{ij}|$ and $\|X\|_\infty = \max_{ij} |X_{ij}|$ then the function f and its conjugate can be represented as

$$f(X) = \|P(A)\|_1, \quad f^*(Y) = \begin{cases} 0, & P(Y) = 0, \ \ \|P^c(Y)\|_\infty < 1 \\ \infty, & \text{otherwise.} \end{cases} \tag{9.27}$$

We note that f is related to the 1-norm of a matrix whose conjugate function is the indicator function associated with the set of the unit ∞-norm ball. The condition in the domain of f^* is almost like the unit ∞-norm ball with an additional condition on the entries of Y belonging to I_A. □

9.6 Projections

In this section, we give the closed form solution to the problem of projecting a matrix A on the a convex set, *i.e.*, we find the matrix lying on $\mathcal{C}$ that is nearest to A measured by the Euclidean norm. The mathematical formulation of this problem is given by

$$\Pi_{\mathcal{C}}(A) \;=\; \underset{X\in\mathcal{C}}{\operatorname{argmin}}\; \frac{1}{2}\|X-A\|_F^2,$$

where $\Pi_{\mathcal{C}}(A)$ denotes the projection of A on $\mathcal{C}$. For some convex sets, the projection can be given in an analytical form as follows [21, 33].

Positive definite cone

If $\mathcal{C} = \mathbf{S}_+^n$ then

$$\Pi_{\mathcal{C}}(A) = \sum_i (\lambda_i)^+ u_i u_i^T \tag{9.28}$$

where $\sum_{i=1}^n \lambda_i u_i u_i^T$ is the eigenvalue decomposition of A, *i.e.*, we discard all the modes corresponding to the negative eigenvalues.

Positive definite cone with an upper bound

For $\mathcal{C} = \{X \in \mathbf{S}^n \mid 0 \preceq X \preceq \alpha I\}$ where $\alpha > 0$ is a parameter, we have

$$\Pi_{\mathcal{C}}(A) = \sum_i \min(\max(\lambda_i, 0), \alpha)\; u_i u_i^T \tag{9.29}$$

where $\sum_{i=1}^n \lambda_i u_i u_i^T$ is the eigenvalue decomposition of A, *i.e.*, we project each eigenvalue of A onto the interval $0 \le \lambda_i \le \alpha$ and recompose A with the corresponding eigenvectors.

Affine subspace

For $\mathcal{C} = \{X \in \mathbf{R}^{n\times n} \mid X_{ij} = b_{ij} \quad \text{for } (i,j) \in I\}$, where the index set I and the values $b_{ij} \in \mathbf{R}$ are given, obviously we have

$$\Pi_{\mathcal{C}}(A) = \begin{cases} A_{ij}, & (i,j) \notin I, \\ b_{ij}, & (i,j) \in I. \end{cases} \tag{9.30}$$

9.7 MATLAB code of ADMM

For this part, we provide MATLAB codes of functions: *my_primal_CVX*() for solving the primal convex SEM formulation and *my_sparse_SEM*() for solving the sparse SEM formulation. These functions were used in our simulation process in section 7.3 and, in section 7.4, the performance of ADMM for solving our estimation formulations were also evaluated by the use of these functions. Each of them contains subfunctions for solving the projection problem so that we would like to firstly introduce these subfunctions which are described as follows.

- *calpdfA*. This function is to find the optimal solution of the projection problem on positive definite cone set which is a problem (6.8) or (6.12) and its solution can be computed as explained in Appendix 9.6. In the detail, this function performs the eigenvalue decomposition of an input matrix and discards all modes corresponding to its negative eigenvalues. Codes of this function are explained below.

```
function[X] = calpdfA(A)
% example : X = calpdfA(A)
% calpdfA calculates the positive part of A
% A = UDU' = [U1 U2]*diag(D1,D2)*[U1'; U2']
% where D1 contains positive eigenvalues of A and
%       D2 contains negative eigenvalues of A

%eigen decomp
[U,D] = eig(A);
d = diag(D);
I1 = find(d >= 0);

U1 = U(:,I1);

X = U1*D(I1,I1)*U1';
```

- *projupperbnd*. This function is to find the optimal solution of the projection problem on positive definite cone set with an upper bound, which is the problem (6.13). Practically, this function requires an $2n \times 2n$ input matrix and the block $(2, 2)$ of this matrix is performed the eigenvalue decomposition as detail explained in Appendix 9.6. We provide MATLAB codes of this function as follows.

```
function[X] = projupperbnd(A,alpha)
% example : X = projupperbnd(A,alpha)
%
% minimize || X - A ||_F
% subject to 0 <= X4 <= alpha*I
% where X = [X1 X2' ; X2 X4]; alpha is a positive scalar
%
% A is 2n x 2n
% Suppose A(2,2) = UDU' %% block (2,2) of A
% and D= diag(d1,d2,...,dn)
% We project the eigenvalues of block (2,2) of A into the
% interval (0,alpha) and X4 is obtained by UP(D)U' where
% P(D) is the projection of eigenvalues
% on the interval (0,alpha)
```

```
    dim = size(A,1);
    n = dim/2;
    TMP = A(n+1:2*n,n+1:2*n);

    [U,D] = eig(TMP);
    d = diag(D);
    PD = diag(min(max(d,0),alpha)); % threhold eigen of A in range
    X4 = U*PD*U';

    X = [A(1:n,1:n) A(1:n,n+1:2*n) ; A(n+1:2*n,1:n) X4];
```

9.7.1 MATLAB code of ADMM for solving the primal convex SEM

This part provides MATLAB codes of a *main* function used for solving the primal convex SEM formulation (6.3). The numerical method used in the function is based on ADMM algorithm corresponding the detail described in section 6.1. These codes are explained step by step below.

```
function [myX,myval,IterUsed] = my_primal_CVX(S,indA,alpha,Ite)
% this function is to solve primal convex SEM based on ADMM algorithm
% it requires input as follows.
%   1) S : sample covariance matrix
%   2) indA : index (i,j) that A_ij is zero (index of I_A)
%   3) alpha : nomally we use alpha = min(eig(S))
%   4) Ite : the number of maximum iteration
%
% and it returns output as follows
%   1) myX : our solution X = [X1 X2 ; X2' X4];
%   2) myval : the value of cost objective
%   3) IterUsed : the number of iteration used in the process
% example :
% [myX,myval,IterUsed] = my_primal_CVX(S,indA,alpha,Ite);

n = length(S(1,:));  %computing n (# observed variables)
k=1;     %iteration count
E_abs = 1e-6;    %absoluite tolerance
E_rel = 1e-6;    %relative tolerance

inddiag = find(eye(n));
indoffA = setdiff(indA,inddiag);

%initial : setting U,V,W in its domain
roh = 1+0.1*randn(); %penalty parameter can be changed
```

```
% U >= 0
U = eye(2*n);
U1 = U(1:n,1:n);
U2 = U(n+1:2*n,1:n);
U4 = U(n+1:2*n,n+1:2*n);

% P(V2) = I and V4 = alpha*eye(n)
V1 = zeros(n,n);    V2 = eye(n);    V4 = alpha*eye(n);
V = [V1 V2'; V2 V4];

%initial : setting the dual variables, Y1 and Y2
Y1 = eye(2*n);
Y1_11 = Y1(1:n,1:n); %block(1,1) of Y1
Y1_21 = Y1(n+1:2*n,1:n);     %block(2,1) of Y1
Y1_22 = Y1(n+1:2*n,n+1:2*n); %block(2,2) of Y1

Y2 = eye(2*n);
Y2_11 = Y2(1:n,1:n);
Y2_21 = Y2(n+1:2*n,1:n);
Y2_22 = Y2(n+1:2*n,n+1:2*n);

%start algorithm
while k<=Ite
    %update X
    if k>1
        Xold = X;   %keep X before update
    end
    X2 = (1/2)*((U2+V2)-(1/roh)*(Y1_21+Y2_21));
    X4 = (1/2)*((U4+V4)-(1/roh)*(Y1_22+Y2_22));
    [eigen_vec eigen_val] = eig(roh*(U1+V1)-(S+Y1_11+Y2_11));
    r = diag(eigen_val);
    X1_ii = (r + sqrt(r.^2+(8*roh)))/(4*roh);
    X1 = eigen_vec*diag(X1_ii)*eigen_vec';
    X = [X1 X2';X2 X4];

    %update U
    Uold = U;   %keep U before update
    U = X+(1/roh)*Y1;
    U = calpdfA(U);
    U1 = U(1:n,1:n);
    U2 = U(n+1:2*n,1:n);
    U4 = U(n+1:2*n,n+1:2*n);
```

```
    %update V
    Vold = V;   %keep V before update
    TMP_V = X+(1/roh)*Y2;
    V1 = TMP_V(1:n,1:n);
    V2 = TMP_V(n+1:2*n,1:n);  V2(inddiag) = 1; V2(indoffA) = 0;
    V4 = alpha*eye(n);
    V = [V1 V2'; V2 V4];

    %update dual variables, Y1 and Y2
    Y1 = Y1+roh*(X-U);
    Y1_11 = Y1(1:n,1:n);
    Y1_21 = Y1(n+1:2*n,1:n);
    Y1_22 = Y1(n+1:2*n,n+1:2*n);

    Y2 = Y2+roh*(X-V);
    Y2_11 = Y2(1:n,1:n);
    Y2_21 = Y2(n+1:2*n,1:n);
    Y2_22 = Y2(n+1:2*n,n+1:2*n);

    %stopping criterion
    if k > 1
        %M1 is used for calculating the stopping criterion
        M1 = [norm(X,'fro'),norm(U,'fro'),norm(V,'fro')];
        %M2 is used for calculating the stopping criterion
        M2 = [norm(Y1,'fro'),norm(Y2,'fro')];

        E_pri = sqrt(4*n^2)*E_abs + E_rel*max(M1);
        E_dual = sqrt(4*n^2)*E_abs + E_rel*max(M2);
        r = norm([X-U;X-V]);
        s = roh*norm([X-Xold;U-Uold;V-Vold]);

        %break function when stopping criterion satisfied
        if (r <= E_pri) && (s <= E_dual )
            myX = X;
            myval = -log_det(X1)+trace(S*X1);
            IterUsed = k;
            break;
        end
    end
    k=k+1;
end
```

9.7.2 MATLAB code of ADMM for solving the sparse SEM

We provide MATLAB codes of a *main* function used for solving the sparse SEM formulation (6.9) in this section. The numerical method used in the function is based on ADMM algorithm corresponding the detail described in section 6.2. These codes are given below.

```
function [X,A,myval,IterUsed] = my_sparse_SEM(S,indA,alpha,gmma,Ite)
% this function is to solve the sparse SEM based on ADMM algorithm
% it requires input as follows.
% 1) S : sample covariance matrix
% 2) indA : index (i,j) that A_ij is zero (index of I_A)
% 3) alpha : we choose alpha small enough to get low rank sol
% 4) gmma or gamma : regularization parameters
% 5) Ite : the number of maximum iteration
%
% and it returns output as follows
% 1) X : our solution X = [X1 X2 ; X2 X4];
% 2) A : our estimated path matrix
% 2) myval : the value of cost objective
% 3) IterUsed : the number of iteration used in the process
% example :
% [X,A,myval,IterUsed] = my_sparse_SEM(S,indA,alpha,gmma,Ite);

n = length(S(1,:));     %computing n (# observed variables)
j=1;                    %iteration count
E_abs = 1e-6;           %absolute tolerance
E_rel = 1e-6;           %relative tolerance

R = ones(n,n);
R(indA) = 0;
indAc = find(R==1);     %index where A_{ij} = 0

% initial : setting Z,U and V in its domain
roh = max(0.15*n,50); % roh : penalty parameter

%Z s.t. P(Z) = 0;
Z = zeros(n,n);

%U s.t. U >= 0
U = eye(2*n);

%V s.t. 0 <= V4 <= alpha*I
V1 = zeros(n,n); V2 = zeros(n,n); V4 = alpha*eye(n);
V = [V1 V2'; V2 V4];
```

```
%initial : setting dual variables, Y1,Y2 and Y3
Y1 = eye(n);
Y2 = eye(2*n);
Y3 = eye(2*n);

%start algorithm
while j<Ite
    %update X
    if j > 1
        Xold = X; %keep X before update
    end
    M = (1/2)*(U+V) - (1/(2*roh))*(Y2+Y3);
    M1 = M(1:n,1:n); M2 = M(n+1:2*n,1:n); M4 = M(n+1:2*n,n+1:2*n);
    H = eye(n) - Z - (1/roh)*Y1;
    X2 = (1/5)*(H+4*M2);
    X4 = M4;
   [eigen_vec, eigen_val] = eig(2*roh*M1 - S);
    r = diag(eigen_val);
    X1_ii = (r + sqrt(r.^2+(8*roh)))/(4*roh);
    X1 = eigen_vec*diag(X1_ii)*eigen_vec';
    X = [X1 X2'; X2 X4];

    %update Z
    Zold = Z; %keep Z before update
    %performing elementwise soft thresholding in Eq 6.11
    k = (2*gmma)/roh;
    TMP_A1 = (eye(n)-X2-(1/roh)*Y1)-k; TMP_IND = find(TMP_A1 <= 0);
    TMP_A1(TMP_IND) = 0;
    TMP_A2 = -(eye(n)-X2-(1/roh)*Y1)-k; TMP_IND = find(TMP_A2 <= 0);
    TMP_A2(TMP_IND) = 0;
    Z = TMP_A1 - TMP_A2; Z(indA) = 0;

    %update U
    Uold = U; %keep U before update
    U = X+(1/roh)*Y2;
    U = calpdfA(U);

    %update V
    Vold = V; %keep V before update
    V = X+(1/roh)*Y3;
    V = projupperbnd(V,alpha);
```

```
    %update dual variables, Y1, Y2 and Y3
    Y1 = Y1+roh*(X2+Z-eye(n));
    Y2 = Y2+roh*(X-U);
    Y3 = Y3+roh*(X-V);

    %stopping criterion
    if j > 1
        %M1 is used for calculating the stopping criterion
        M1 = [norm(X,'fro'),norm(U,'fro'),norm(V,'fro'),norm(Z,'fro')];
        %M2 is used for calculating the stopping criterion
        M2 = [norm(Y1,'fro'),norm(Y2,'fro'),norm(Y3,'fro')];

        E_pri = sqrt(4*n^2)*E_abs + E_rel*max(M1);
        E_dual = sqrt(4*n^2)*E_abs + E_rel*max(M2);
        r1 = norm([X-U;X-V]);
        r2 = norm(X2 + Z - eye(n));
        s1 = roh*norm([X-Xold;U-Uold;V-Vold]);
        s2 = roh*norm(Z-Zold);

        %break program
        if (r1 <= E_pri && r2 <= E_pri) && ...
            (s1 <= E_dual && s2 <= E_dual)
            X = X;
            A = Z;
            myval = -log_det(X1)+trace(S*X1)...
                    +2*gmma*sum(abs(Z(indAc)));
            IterUsed = j;
        break;
        end
    end
    j=j+1;
end
```

Biography

Anupon Pruttiakaravanich was born in Bangkok, Thailand, in 1988. He received his Bachelor's degree with the second honor in electrical engineering from King Mongkut's University of Technology North Bangkok (KMUTNB), in 2011. He had worked as a process engineer at Microchip Technology (Thailand) Co. Ltd for 3 years then he decided to study in higher level. He was admitted to the Master program in electrical engineering at Chulalongkorn University, Thailand in 2014 and he has conducted his graduate study with the Control Systems Research Laboratory, Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University. His research interests include system estimation and system identification in convex optimization framework.

List of Publications

1. A. Pruttiakaravanich and J. Songsiri, "A Review on Exploring Brain Networks from fMRI Data," *Engineering Journal*, Vol.20, No.3, 2016.

2. A. Pruttiakaravanich and J. Songsiri, "A Convex Formulation for Path Analysis in Structural Equation Modeling," *Proc. of SICE Annual Conference (SICE 2016)*, Japan, 2016.