A Modified Decomposition Covariance Matrix Estimation for Undirected Gaussian Graphical Model

Ridawarni P., Yenny Hermiana A., Saprilina G.
Graduate School of Mathematics,
University of Sumatera Utara
Medan, Indonesia

Abstract - A covariance matrix is an undirected graph that associated with a multivariate probability of a given random vector where each vertex represents the different components of the random vector. Graphical model are framework for representing and conditional independence structures with distribution using graph $G$. This paper discussed a distribution estimation in determining decomposable covariance matrix in an undirected Gauss graphical model related to sparsity of invers covariance (concentration matrix). It showed decomposable covariance estimation with lower computational complexity. The result showed the correlation each different components in a given random vector that determined from decomposition covariance matrix estimation.

Keywords - conditional independence, covariance decomposition, Gauss graphical model, concentration graph

I. INTRODUCTION

Graphical models are a framework in determine a conditional independence structures within distributions that represented by a graph. In representing a distribution, undirected graph is used as a common model to describe a distribution problem in high dimension. Some previous researches that related to undirected graph model has been successfully applied to determine a conditional independence structure in a multivariate distribution, and extended with a computation method using a representation by a graph, especially for dimension enhancement and complexity problem. The results showed a sum of weighted path of all available paths that connected the two variables in an undirected independence graph [11].

An estimation of decomposition covariance matrix is the main focus topic in this paper. The estimation is focused on an undirected Gaussian graphical model of two random variables that gives correlation of each vertex on a Gaussian graph as a result.

II. REVIEWS

The Gaussian distribution is also referred to as the normal distribution or the bell curve distribution for its bell-shaped density curve. The formula for a $d$-dimension Gaussian probability distribution is

$$p(x) = \frac{1}{(2\pi)^{d/2}} \left| \Sigma \right|^{1/2} \exp \left( -\frac{(x-\mu)^T \Sigma^{-1} (x-\mu)}{2} \right)$$ (2.1)

where $x$ is a $d$-element column vector of variables along each dimension, $\mu$ is the mean vector, calculated by

$$\mu = E[x] = \int x \rho(x) \, dx$$ (2.2)

and $\Sigma$ is the $d \times d$ covariance matrix, calculated by

$$\Sigma = E[(x-\mu)(x-\mu^T)] = \int (x-\mu)(x-\mu^T) p(x) \, dx$$ (2.3)

with the following form
The covariance matrix is always symmetric and positive semidefinite, where positive semidefinite means that for all non-zero $x \in R^d, x^T \Sigma x \geq 0$. We normally only deal with covariance matrices that are positive definite where for all non-zero $x \in R^d, x^T \Sigma x > 0$, such the determinant $|\Sigma|$ will be strictly positive. The diagonal elements, $\sigma_{ii}$, are the variances of the respective $x_i$, i.e., $\sigma_i^2$, and the off-diagonal elements, $\sigma_{ij}$, are the covariances of $x_i$ and $x_j$. If the variables along each dimension is statistically independent, then $\sigma_{ij} = 0$, and we would have a diagonal covariance matrix

$$
\begin{bmatrix}
\sigma_{11} & \cdots & \sigma_{1d} \\
\sigma_{21} & \cdots & \sigma_{2d} \\
\vdots & \cdots & \vdots \\
\sigma_{d1} & \cdots & \sigma_{dd}
\end{bmatrix}
$$

(2.4)

The covariance matrix is always symmetric and positive semidefinite, where positive semidefinite means that for all non-zero $x \in R^d, x^T \Sigma x \geq 0$. We normally only deal with covariance matrices that are positive definite where for all non-zero $x \in R^d, x^T \Sigma x > 0$, such that the determinant $|\Sigma|$ will be strictly positive. The diagonal elements, $\sigma_{ii}$, are the variances of the respective $x_i$, i.e., $\sigma_i^2$, and the off-diagonal elements, $\sigma_{ij}$, are the covariances of $x_i$ and $x_j$. If the variables along each dimension is statistically independent, then $\sigma_{ij} = 0$, and we would have a diagonal covariance matrix

$$
\begin{bmatrix}
\sigma_1^2 & 0 & \cdots & 0 \\
0 & \sigma_2^2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \sigma_d^2
\end{bmatrix}
$$

(2.5)

If the covariances along each dimension is the same, then we will have an identify matrix multiplied by a scalar,

$$
\sigma^2 I
$$

by the Eq. (2.6), the determinant of $\Sigma$ becomes

$$
|\Sigma| = \sigma^{2d}
$$

(2.7)

and the inverse of $\Sigma$ becomes

$$
\Sigma^{-1} = \begin{bmatrix}
\frac{1}{\sigma_1^2} & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & \frac{1}{\sigma_d^2}
\end{bmatrix}
$$

(2.8)

For 2-d Gaussian where $d = 2$, $x = [x_1 \ x_2]^T$, $|\Sigma| = \sigma^4$, the formulation becomes

$$
p(x_1, x_2) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x_1^2 + x_2^2}{2\sigma^2}\right)
$$

(2.9)

then we often denote a Gaussian distribution of Eq. (2.1) as $p(x) \square N(\mu, \Sigma)$.

An undirected graph $G = (V, E)$ is a set of nodes $V = \{1, \ldots, |V|\}$ connected by undirected edges $E = \{(i_1, j_1), \ldots, (i_{|E|}, j_{|E|})\}$, where each node is connected to itself, i.e., $(i, i) \in E$ for all $i \in V$. Let $x = [x_1, \ldots, x_p]^T$ be a zero random vector of length $p = |V|$ whose elements are indexed by the nodes in $V$. The vector $x$ satisfies the Markov property with respect to $G$, if for any pair of nonadjacent nodes the corresponding pair of elements in $x$ are conditionally independent of the remaining elements, i.e., $x_i$ and $x_j$ are conditionally independent of $x_r$ for any $(i, j) \notin E$ and $r = |V \setminus i, j|$ where

$$
p(x_i, x_j | x_r) = p(x_i | x_r) p(x_j | x_r)
$$

(2.10)

Therefore, the joint distribution satisfies the following factorization:

$$
p(x_i, x_j, x_r) = \frac{p(x_i, x_j) p(x_j, x_r)}{p(x_r)}
$$

(2.11)

In the Gaussian setting, this factorization leads to sparsity in the concentration (inverse covariance) matrix. The multivariate Gaussian distribution is defined as

$$
p(x; K) = c |K|^{-\frac{1}{2}} e^{-\frac{1}{2}x^T Kx}
$$

(2.12)

where $c^*$ in an appropriate constant and the marginal concentration matrix is

$$
\bar{K}_r = (K^{-1})_{r,r}^{-1}
$$

(2.13)

Together with Eq. (2.11) this implies that

$$
\bar{K}_r = (K^{-1})_{r,r}^{-1}; K = [\bar{K}_\rho]_\rho + [\bar{K}_\rho]_\rho^T - [\bar{K}_\rho]^T
$$

(2.14)

$$
|K| = \left| \bar{K}_\rho \right| \left| \bar{K}_\rho \right|
$$

(2.15)

where $\bar{K}_\rho$, $\bar{K}_\rho^T$ and $\bar{K}_\rho$ are the marginal concentrations of $\{x_i, x_r\}$, $\{x_j, x_r\}$ and $\{x_r\}$, respectively, and are defined in a similar manner to Eq. (2.13). All the matrices in the right hand side of Eq. (2.14) have a zero value in the $(i, j)$th position, and therefore

$$
[K]_{i,j} = 0 \text{ for all } (i, j) \notin E
$$

This property is the core of Gaussian graphical models; the concentration matrix $K$ has a sparsity pattern which represents the topology of the conditional independence graph.

Decomposable models are a special type of graphical model in which the conditional independence graphs satisfy an appealing structure. A decomposable graph can be divided
III. DECOMPOSITION COVARIANCE MATRIX

In our estimation, we use some notations in determine covariance matrix in Gaussian graphical model. Assume \( X = X_{V_p} \) where \( V_p = \{1, \ldots, p\} \) is a random vector with normal multivariate distribution with \( p \)-dimension, \( N_p(0, K^{-1}) \). Set a graph \( G = (V_p, E) \) where for each vertex \( i \in V \) is pair to a variable \( X_i \) and \( E \subset V_p \times V_p \) which is an undirected graph. Set the definition that \((i, j) \in E\) if and only if \((j, i) \in E\). A Gaussian graphical model with conditional independence graph is determined with the limit of diagonal elements of \( K \) that not pair to any edge in \( G \). If \((i, h) \notin E\), then \( X_i \) and \( X_j \) is conditionally independence of a given random variables. Concentration matrix \( K = (K_{ij})_{i,j \in V_p} \) is a limit to a symmetric positive definite matrix with diagonal entry \( K_{ij} = 0 \) for all \((i, j) \notin E\). Using G-Wishart distribution, \( \text{Wish}_G(\delta, D) \) with density

\[
p(K|G, \delta, D) = \frac{1}{\Gamma_G(\delta/2)} |K|^{(\delta/2) - 1} \exp\left(-\frac{1}{2} \text{tr}(K)\right) \quad (3.1)
\]

based on Lebesgue estimation (see [12]; [13]; [14]). If \( G \) is a complete graph \( E = (V_p \times V_p) \), then \( \text{Wish}_G(\delta, D) \) is a G-Wishart distribution \( \text{Wish}_G(\delta, D) \). We then using Cholesky decomposition for matrix \( K \) with \( K \in P_G \) is \( K \in P_G \) where \( Q = (q_{ij})_{1 \leq i \leq j \leq p} \) and \( \psi = (\psi_{ij})_{1 \leq i \leq j \leq p} \) is an upper triangle where \( D^{-1} = Q^TQ \) is the decomposition Cholesky of \( D^{-1} \).

Estimation of covariance matrix is a basic problem of multivariate statistical that related to signal processing, financial mathematics, pattern recognition and convex geometry computation. Set a sample of \( n \) points that independent, \( X_1, \ldots, X_n \), from distribution. We then have a sample of covariance matrix

\[
\Sigma_n = \frac{1}{n} \sum_{i=1}^{n} X_i X_i^T = \frac{1}{n} A^T A \quad (3.1)
\]

with \( \Sigma_n \) is a random matrix. Then we estimated the covariance matrix \( \Sigma \) with rate of accuracy, \( \varepsilon = 0.01 \) that represented by a norm operation as follows.

\[
\|\Sigma_n - \Sigma\| \leq \varepsilon \quad (3.2)
\]

Assume \( X = (X_1, X_2, \ldots, X_n) \) of a multivariate Gaussian distribution \( N_G(0, \Sigma) \) where \( \Sigma \) is a regular matrix. Then we determined likelihood function

\[
L(K) = \frac{(\det K)^{n/2} e^{-\text{tr}(Kw)}}{2} \quad (3.3)
\]

where \( W = \sum_{i=1}^{n} X_i^T(X_i^T)^T \) as a result of sum and multiple matrix. Thus, the likelihood equation of Gaussian matrix can be formulated as

\[
\hat{K}^{-1} = \frac{n \sum}{w} \quad (3.5)
\]

\[
\log L(K) = \frac{n}{2} \log(\det K) - \frac{\text{tr}(Kw)}{2} \quad (3.6)
\]

\[
\frac{\partial}{\partial k_{ij}} \log(\det K) = \frac{w_{ij}}{n} \quad (3.7)
\]

\[
\frac{\partial}{\partial k_{ij}} \log(\det K) = K^{-1} \quad (3.8)
\]

IV. DECOMPOSITION COVARIANCE MATRIX FOR UNDIRECTED GAUSSIAN GRAPH

Adopted by a basic structure of variance component and time series problem, we suggest definition of linear covariance formula model that represented as

\[
\Sigma = a_1 U_1 + \ldots + a_q U_q \quad (4.1)
\]

where \( U_i \) are symmetric matrices and \( a_i \) is an unknown parameter that supposed to be a requirement so that the matrix is positive definite. Eq. (4.1) represents a common formula of time-series covariance model, mixed-linear and graph model. Specifically, high dimension \( q \) for any covariance matrix can be denoted as

\[
\Sigma = (\sigma_{ij}) = \sum_{i=1}^{p} \sum_{j=1}^{q} \sigma_{ij} U_{ij} \quad (4.2)
\]

with \( U_{ij} \) is matrix with dimension \( p \times p \) with element 1 at \((i,j)\) and 0 otherwise. For each column and row, we can set variance matrix with these following steps.

Step 1. Set matrix \( X \) as a deviation for \( x \) where \( x = X - 11X(\frac{1}{n}) \).

Step 2. Calculate \( x'x \) as a result of sum and multiple matrix with dimension \( k \times k \) at matrix \( x \).

Step 3. Each element of matrix \( x'x \) is divided by \( n \). Thus, we determined variance-covariance matrix of matrix \( x \),

\[
V = x'x(\frac{1}{n}) \quad (3.9)
\]

where

\[
\begin{align*}
1 & : \text{column vector with element 1 of dimension } n \times 1 \\
x & : \text{deviation matrix of dimension } m \times n \\
X & : \text{data matrix of dimension } m \times n \\
V & : \text{covariance matrix of dimension } m \times n \\
x'x : \text{result of sum and multiple matrix} \\
n & : \text{number of trial of matrix } X \\
\end{align*}
\]
Thus, we then determined a decomposition covariance matrix that showed correlation partial of the two random variables by

$$\rho_{ij} = \frac{k_{ij}}{\sqrt{k_{ii}k_{jj}}}$$  \hspace{1cm} (4.3)

V. RESULTS

A. Matrix m × m

For illustration, we use a given matrix data of dimension $m \times m$. Assume that there exist matrix $A_{3\times3}$ as follows.

$$A_{3\times3} = \begin{pmatrix} 1 & 4 & 1 \\ 2 & 1 & 4 \\ 3 & 3 & 1 \end{pmatrix}$$

Then, covariance matrix of matrix $A$ is determined by the definition $Cov(A) = A - AA'$. Thus,

$$Cov(A) = \begin{pmatrix} -5 & -4 & -5 \\ -4 & -7 & -2 \\ -3 & -5 & -5 \end{pmatrix}$$

Because $n = 3$, for each element of covariance matrix we have

$$Cov(A) = \begin{pmatrix} 0.1 & 0.08 & 0.004 \\ 0.1 & 0.02 & 0.014 \\ 0.0 & 0.12 & 0.026 \end{pmatrix}$$

with the decomposition of inverse matrix is

$$inv(A) = \begin{pmatrix} -1.172 & -0.234 & 1.266 \\ 0.656 & -0.469 & -0.469 \\ 0.047 & 0.609 & -0.890 \end{pmatrix}$$

and for each correlation of the two random variables, we determined

$$\rho_{123} = \frac{0.2344}{\sqrt{(-1.1719)(-0.4688)}} = \frac{0.2344}{0.7412} = 0.3162$$

$$\rho_{132} = \frac{-0.0469}{\sqrt{(-1.1719)(-0.8906)}} = \frac{-0.0469}{1.0216} = -0.0459$$

$$\rho_{213} = \frac{0.4788}{\sqrt{(-0.4688)(-0.8906)}} = \frac{0.4788}{0.6461} = 0.7255$$

B. Matrix m × n

Assume that there exist matrix

$$A_{5\times3} = \begin{pmatrix} 4.0 & 2.0 & 0.60 \\ 4.2 & 2.1 & 0.59 \\ 3.9 & 2.0 & 0.58 \\ 4.3 & 2.1 & 0.62 \\ 4.1 & 2.2 & 0.63 \end{pmatrix}$$

which the covariance matrix of matrix $A$ is determined by the definition $Cov(A) = A - AA'$. Thus,

$$Cov(A) = \begin{pmatrix} -0.1 & -0.08 & -0.004 \\ 0.1 & 0.02 & -0.014 \\ -0.2 & -0.08 & -0.024 \\ 0.2 & 0.02 & 0.016 \\ 0.0 & 0.12 & 0.026 \end{pmatrix}$$

Because $n = 5$, for each element of covariance matrix we have

$$Cov(A) = \begin{pmatrix} -0.1 & -0.08 & -0.004 \\ 0.1 & 0.02 & -0.014 \\ -0.2 & -0.08 & -0.024 \\ 0.2 & 0.02 & 0.016 \\ 0.0 & 0.12 & 0.026 \end{pmatrix} \begin{pmatrix} 1 \\ 5 \end{pmatrix}$$

$$= \begin{pmatrix} 0.02 & 0.006 & 0.0014 \\ 0.006 & 0.0056 & 0.0011 \\ 0.0014 & 0.0011 & 0.0003 \end{pmatrix}$$

with the decomposition of inverse matrix is

$$inv(A) = \begin{pmatrix} 76.1 & 55.3 & 136.2 \\ 20.8 & 492.8 & 1322.1 \\ -136.2 & 1322.1 & 7612.2 \end{pmatrix}$$

and for each correlation of the two random variables, we determined

$$\rho_{123} = \frac{55.3}{\sqrt{(76.1)(492.8)}} = \frac{55.3}{193.65} = 0.285$$

$$\rho_{132} = \frac{55.3}{\sqrt{7612.2}} = \frac{55.3}{761.09} = 0.072$$

$$\rho_{213} = \frac{1322.1}{\sqrt{(492.8)(7612.2)}} = \frac{1322.1}{1936.825} = 0.682$$
VI. CONCLUSIONS

In this paper, we studied the decomposition of covariance for undirected Gaussian graph. Estimation of covariance in a Gaussian distribution is a common problem in statistic. Gaussian graphical model is a method that can be used to represent the structure are independent among different independence random variables in a graph. This paper has examined the results of the covariance estimation for undirected Gaussian by using a likelihood function in order to obtain the concentration of each random variables. The results of the decomposition matrix is decomposed and can be used in solving problems that related to signal transmission, pattern recognition and other concentrations matrix problem.

REFERENCES


