

## Topic 16: Graphical Network Inference

by Sai Zhang

**Key points:**

**Disclaimer:** The note is built on Prof. *Jinchi Lv*'s lectures of the course at USC, DSO 607, High-Dimensional Statistics and Big Data Problems.

**16.1 Motivation**

Consider a classic question: For  $n$  observations of dimension  $p$ , how can we capture the statistical relationships between the variables of interest? Consider the example of the multivariate Gaussian distribution:

**Example 16.1.1: Multivariate Gaussian Distribution**

Suppose we have  $n$  observations of dimension  $p$ ,  $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ . let  $\mathbf{S}$  be the empirical covariance matrix. Then the probability density

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{p/2} \det(\Sigma)^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mu)' \Sigma^{-1} (\mathbf{x} - \mu) \right\}$$

define the **inverse covariance matrix** or **precision matrix** as  $\Omega = \Sigma^{-1}$ , then we have

$$f_{\mu, \Omega} = \exp \left\{ \mu' \Omega \mathbf{x} - \left\langle \Omega, \frac{1}{2} \mathbf{x} \mathbf{x}' \right\rangle - \frac{p}{2} \log(2\pi) + \frac{1}{2} \log \det(\Omega) - \frac{1}{2} \mu' \Omega \mu \right\}$$

In this example, we know that **every** multivariate Gaussian distribution can be represented by a pairwise **Gaussian Markov Random Field (GMRF)**, which an **undirected graph**  $G = (V, E)$

- representing the collection of variables  $\mathbf{x}$  by a vertex set  $\mathcal{V} = \{1, \dots, p\}$
- encoding correlations between variables by a set of edges  $\mathcal{E} = \{(i, j) \in \mathcal{V} \mid i \neq j, \Omega_{ij} \neq 0\}$

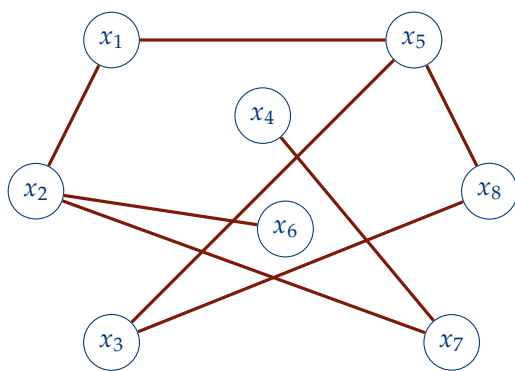
**What's the goal?** We want to estimate a **sparse** graph structure given  $n \ll p$  i.i.d. observations. But what does sparsity means in this context? A sparse graph is **equivalent** to a sparse precision matrix: the precision matrix should have many 0s.

**Sparse precision matrix** for the Gaussian vector mentioned above  $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ , we have  $\forall u, v$

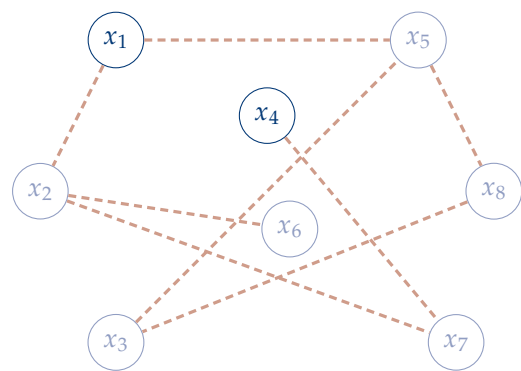
$$x_u \perp x_v \mid \mathbf{x}_{\mathcal{V} \setminus \{u, v\}} \Leftrightarrow \Theta_{u, v} = 0$$

that is, sparsity of the precision matrix is equivalent to **conditional independence**<sup>1</sup>. Consider a graph

<sup>1</sup>Meanwhile, for independence:  $\Sigma_{u, v} = 0 \Leftrightarrow x_u \perp x_v$



$x_1$  and  $x_4$  are connected



$x_1$  and  $x_4$  are NOT connected, conditionally