

Final Project 2

贝叶斯高斯图模型

网络分析最近受到了很多统计学家的关注。一个常用的网络分析模型是高斯图模型 (Gaussian Graphical Model)。在高斯分布的假设下 (即 $X \sim N(\mu, \Sigma), X \in R^p$)，随机向量 X 的两个分量的条件独立性等价于精细矩阵的对应元素为零，即，如果 $\Theta = \Sigma^{-1}$ 为精细矩阵 (协方差矩阵的逆)，则充分必要条件是 $\theta_{ij} = 0$ 。因此，在高斯分布下，网络等价于精细矩阵，这样的网络分析方法一般称为高斯图模型。很多学者都考虑了高斯图模型的估计，最直接的方式是通过极大化带惩罚的似然函数，如 Yuan and Lin (2007) 考虑了极大化带精细矩阵 Lasso 惩罚的 \log 似然函数的方法来估计高斯图模型

$$\frac{n}{2} \log(\det(\Theta)) - \frac{1}{2} \sum_{i=1}^n (X_i - \mu)^T \Theta (X_i - \mu) + \lambda \sum_{i \neq j} |\theta_{ij}|。$$

给定任意参数 $\lambda > 0$ ，我们可以极小化此目标函数，然后利用交叉验证等方法选择合适的参数值 λ 。另一方面， X_j 的条件期望为 $E(X_j | X_{\{1, \dots, p\} \setminus \{j\}}) = \sum_{k \neq j} -(\theta_{jk} / \theta_{kk}) X_k$ ，故高斯图模型亦可以用线性回归的方法估计，例如 Meinshausen and Bühlmann (2016) 通过极小化带 Lasso 惩罚的平方损失函数来估计高斯图模型。

项目目标：

发展高斯图模型的贝叶斯模型和方法。详细说明你的贝叶斯模型及其合理性，发展相应的算法、并用模拟和实际数据分析评价你的方法的表现。此文后列出了一些相关参考文献 (Cheng and Lenkoski 2012, Dobra et al. 2011, Green and Thomas 2013, Mohammadi, A., & Wit, E. C. 2015, Peterson et al. 2015, Wang and Li 2012)，你可以参考这些文章发展高斯图模型的贝叶斯模型。

参考文献

Cheng, Y., & Lenkoski, A. (2012). Hierarchical Gaussian graphical models: Beyond reversible jump. *Electronic Journal of Statistics*, 6, 2309-2331.

Dobra, A., Lenkoski, A., & Rodriguez, A. (2011). Bayesian inference for general Gaussian graphical models with application to multivariate lattice data. *Journal of the American Statistical Association*, 106(496), 1418-1433.

Green, P. J., & Thomas, A. (2013). Sampling decomposable graphs using a Markov chain on junction trees. *Biometrika*, *100*(1), 91-110.

Meinshausen, N., & Bühlmann, P. (2006). High-dimensional graphs and variable selection with the lasso. *The annals of statistics*, *34*(3), 1436-1462.

Mohammadi, A., & Wit, E. C. (2015). Bayesian structure learning in sparse Gaussian graphical models. *Bayesian Analysis*, *10*(1), 109-138.

Peterson, C., Stingo, F. C., & Vannucci, M. (2015). Bayesian inference of multiple Gaussian graphical models. *Journal of the American Statistical Association*, *110*(509), 159-174.

Wang, H., & Li, S. Z. (2012). Efficient Gaussian graphical model determination under G-Wishart prior distributions. *Electronic Journal of Statistics*, *6*, 168-198.

Yuan, M., & Lin, Y. (2007). Model selection and estimation in the Gaussian graphical model. *Biometrika*, *94*(1), 19-35.