OSE SEMINAR 2013

Stratified Gaussian Graphical Models

Henrik Nyman

CENTER OF EXCELLENCE IN
OPTIMIZATION AND SYSTEMS ENGINEERING
ÅBO AKADEMI UNIVERSITY

TURKU, NOVEMBER 15 2013

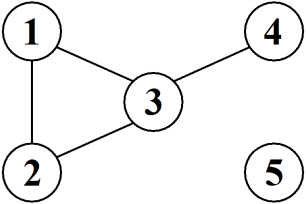
OSE OPTIMIZATION AND SYSTEMS ENGINEERING

Åbo Akademi University

Introduction

- Multivariate Gaussian distribution.
- Gaussian graphical model.
  - Graphical illustration of the dependence structure among the variables in a multivariate Gaussian distribution.
- Multivariate Gaussian distributions constitute a very rigid family of distributions in regard to the dependence structure.
- Introduce context-specific independencies in order to accommodate a more diverse class of distributions and models.

Graphical Models



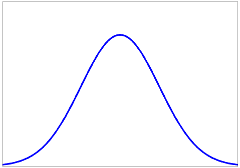
- Marginal independence

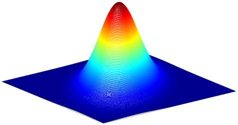
$$X_1 \perp X_5$$

- Conditional independence

$$X_1 \perp X_4 \mid X_3$$

Multivariate Gaussian Distribution



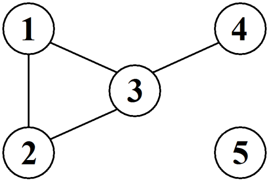


$$X \sim N(\mu, \Sigma)$$

$$f_{\mu,\Sigma}(x) = (2\pi)^{-d/2}|\Sigma|^{-1/2}e^{-1/2(x-\mu)^T\Sigma^{-1}(x-\mu)}$$

$$K = \Sigma^{-1}$$

Graphical Model Meets Multivariate Gaussian



$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} & 0 \\ \sigma_{12} & \sigma_{22} & \sigma_{23} & \sigma_{24} & 0 \\ \sigma_{13} & \sigma_{23} & \sigma_{33} & \sigma_{34} & 0 \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_{44} & 0 \\ 0 & 0 & 0 & 0 & \sigma_{55} \end{pmatrix}$$

$$K = \begin{pmatrix} k_{11} & k_{12} & k_{13} & 0 & 0 \\ k_{12} & k_{22} & k_{23} & 0 & 0 \\ k_{13} & k_{23} & k_{33} & k_{34} & 0 \\ 0 & 0 & k_{34} & k_{44} & 0 \\ 0 & 0 & 0 & 0 & k_{55} \end{pmatrix}$$

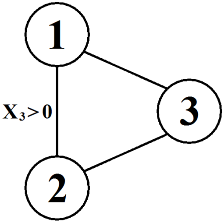
Fitting Σ to Data

- Consider a data matrix $\mathbf{X}$, consisting of n observations (rows) on d variables (columns).
- The maximum likelihood estimates for μ and Σ are calculate

$$\hat{\mu} = \frac{1}{n}\Sigma_{i=1}^{n} x_i \qquad \hat{\Sigma} = \frac{1}{n}\Sigma_{i=1}^{n}(x_i - \hat{\mu})(x_i - \hat{\mu})^T.$$

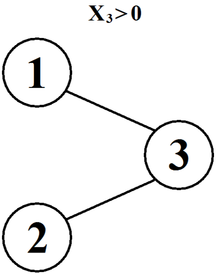
- Using iterative proportional fitting $\hat{\Sigma}$ can be transformed to fit any graph.

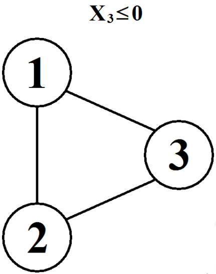
Context-Specific Independence

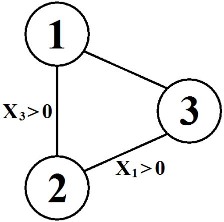


- Conditional independence that holds only in a subset of the outcome space.

$$X_1 \perp X_2 \mid X_3 > 0$$



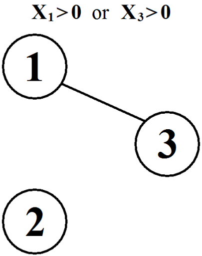


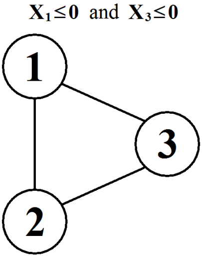


- Two context-specific independencies

$$X_1 \perp X_2 \mid X_3 > 0 \qquad X_3 \perp X_2 \mid X_1 > 0.$$

- For these to hold simultaneously X_2 has to be independent of both X_1 and X_3 once either $X_1 > 0$ or $X_3 > 0$.





- Transforming the strata to a discrete setting will allow for a coherent analysis of the dependence structure.

- This results in a set of conditions on the variables with each condition $c^{(i)}$ associated with a specific dependence structure in the form of a graph $G^{(i)}$.
  - The conditions form a partition of the outcome space of the variables.
  - The condition $c^{(i)}$ can be written in the form $a_j^{(i)} < X_j < b_j^{(i)}$, $j = 1, \ldots, d$.

- Using iterative proportional fitting the graph $G^{(i)}$ induces a specific covariance matrix $\Sigma^{(i)}$.

Density Function for Stratified Gaussian Graphical Models

- The density function of a distribution in a stratified Gaussian graphical model can be written

$$g_{\mu,\Sigma}(x) = \frac{1}{Z}\sum_{i=1}^{\rho} f_{\mu,\Sigma^{(i)}}(x) I_{c^{(i)}}(x)$$

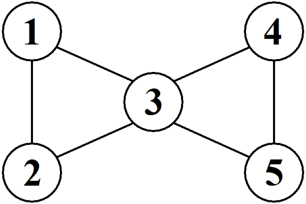
- The normalizing constant Z is calculated as

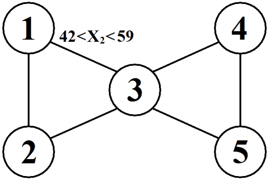
$$Z = \sum_{i=1}^{\rho} \int_{a_1^{(i)}}^{b_1^{(i)}} \dots \int_{a_d^{(i)}}^{b_d^{(i)}} f_{\mu,\Sigma^{(i)}}(x) dx_d \dots dx_1$$

- This distribution belongs to the curved exponential family.
  - Bayesian information criterion can be used to approximate the marginal likelihood of a model.

Math Mark Data

Variable	Label
Mechanics	1
Vectors	2
Algebra	3
Analysis	4
Statistics	5





$$K = \begin{pmatrix} 0.0053 & -0.0025 & -0.0029 & 0.0000 & -0.0001 \\ -0.0025 & 0.0105 & -0.0048 & -0.0008 & -0.0002 \\ -0.0028 & -0.0048 & 0.0273 & -0.0071 & -0.0048 \\ 0.0000 & -0.0008 & -0.0071 & 0.0100 & -0.0020 \\ -0.0001 & -0.0002 & -0.0048 & -0.0020 & 0.0065 \end{pmatrix}$$

Considering only the variables X_1, X_2, and X_3 we get these elements in the precision matrix.

$$k_{13}^{(1)} = -0.0078 \qquad k_{13}^{(2)} = 0.000015$$

References

H. Nyman, J. Pensar, and J. Corander.
Stratified gaussian graphical models.
2013.

H. Nyman, J. Pensar, T. Koski, and J. Corander.
Stratified graphical models - context-specific independence in graphical models.
arXiv:1309.6415 [stat.ML], 2013.

J. Pensar, H. Nyman, T. Koski, and J. Corander.
Labeled directed acyclic graphs: a generalization of context-specific independence in directed graphical models.
arXiv:1310.1187 [stat.ML], 2013.

J. Whittaker.
Graphical models in applied multivariate statistics.
Chichester: Wiley, 1990.

The end of the presentation

Thank you for listening!

The end of the presentation

Thank you for listening!

Questions?