Undirected Gaussian graphical models

Guillaume Obozinski

Swiss Data Science Center

African Masters of Machine Intelligence, 2018-2019, AIMS, Kigali

Multivariate Gaussian distribution

$X \sim \mathcal{N}(\mu, \Sigma)$ with $\mu \in \mathbb{R}^p$, $\Sigma \in \mathbb{R}^{p \times p}$, $\Sigma \succ 0$.

$$p(x, \mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^p |\Sigma|}} \exp\left[-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right]$$

Exponential family form for the multivariate Gaussian

Denoting $\eta = \Sigma^{-1}\mu$ et $\Lambda = \Sigma^{-1}$ we get:

$$\begin{aligned}
(x-\mu)^T \Sigma^{-1} (x-\mu) &= x^T \Sigma^{-1} x - 2\,\mu^T \Sigma^{-1} x + \mu^T \Sigma^{-1} \mu \\
&= x^T \Lambda\, x - 2\,\eta^T x + \eta^T \Lambda^{-1} \eta \\
p\,(x, \mu, \Lambda) &= \exp\left[\eta^T x - \frac{1}{2} x^T \Lambda\, x - A\,(\eta, \Lambda)\right] \\
A\,(\eta, \Lambda) &= \frac{1}{2}\eta^T \Lambda^{-1} \eta + \frac{p}{2} \log 2\pi - \frac{1}{2} \log |\Lambda|
\end{aligned}$$

Canonical parameters: $\theta = \{\Lambda, \eta\}$

- Λ is called the *precision matrix*, $\Lambda \in \mathbb{R}^{p \times p}$
- η is the *loading vector* $\eta \in \mathbb{R}^p$.

Sufficient statistic:

$$\Phi(x) = \begin{pmatrix} x \\ -\frac{1}{2}\,\mathrm{Vec}\left(x x^T\right) \end{pmatrix}$$

Gibbs form of the Gaussian distribution

$$
\begin{aligned}
p(x;\eta,\Lambda) &= \exp\Big(\eta^\top x - \frac{1}{2}x^\top \Lambda x - A(\eta,\Lambda)\Big) \\
&= \frac{1}{Z(\eta)} \prod_{i=1}^{d} \exp(\eta_i x_i - \tfrac{1}{2}\lambda_{ii}x_i^2) \prod_{1\leq i\neq j\leq d} \exp(-\tfrac{1}{2}\lambda_{ij}x_i x_j) \\
&= \frac{1}{Z(\eta)} \prod_{i=1}^{d} \underbrace{\exp(\eta_i x_i - \tfrac{1}{2}\lambda_{ii}x_i^2)}_{\psi_i(x_i)} \prod_{1\leq i< j\leq d} \underbrace{\exp(-\lambda_{ij}x_i x_j)}_{\psi_{ij}(x_i,x_j)}
\end{aligned}
$$

The edges of the graph are $\{(i,j) \mid \lambda_{ij} \neq 0\}$.

When is $\lambda_{ij} = 0$?

Conditional covariance of an edge

Let $I = \{i, j\}$ and $B = V \backslash I$

- $p\left(x_i, x_j | x_B\right)$ is Gaussian with canonical parameters:

$$\eta_{I|B} = \begin{pmatrix} \eta_i - \Lambda_{iB}\, x_B \\ \eta_j - \Lambda_{jB}\, x_B \end{pmatrix} \qquad \text{and} \qquad \Lambda_{II|B} = \Lambda_{II} = \begin{pmatrix} \lambda_{ii} & \lambda_{ij} \\ \lambda_{ji} & \lambda_{jj} \end{pmatrix}$$

Covariance matrix of $X_I | X_B$

$$\mathrm{Cov}\left(X_I | X_B\right) = \Sigma_{II|B} = \Lambda_{II|B}^{-1} = \frac{1}{|\Lambda_{II}|} \begin{pmatrix} \lambda_{jj} & -\lambda_{ij} \\ -\lambda_{ji} & \lambda_{ii} \end{pmatrix}$$

So $\mathrm{corr}\left(x_i, x_j | X_B\right) = \frac{-\lambda_{ij}}{\sqrt{\lambda_{ii}\lambda_{jj}}}$ and

$$\lambda_{ij} = 0 \Leftrightarrow X_i \perp X_j \mid X_B \underset{\text{since Gaussian}}{\Leftrightarrow} X_i \perp\!\!\!\perp X_j \mid X_B$$