CS 583: Probabilistic Graphical Models

Topic: Gaussian Network Models
Chapter: 7

Mustafa Bilgic

http://www.cs.iit.edu/~mbilgic

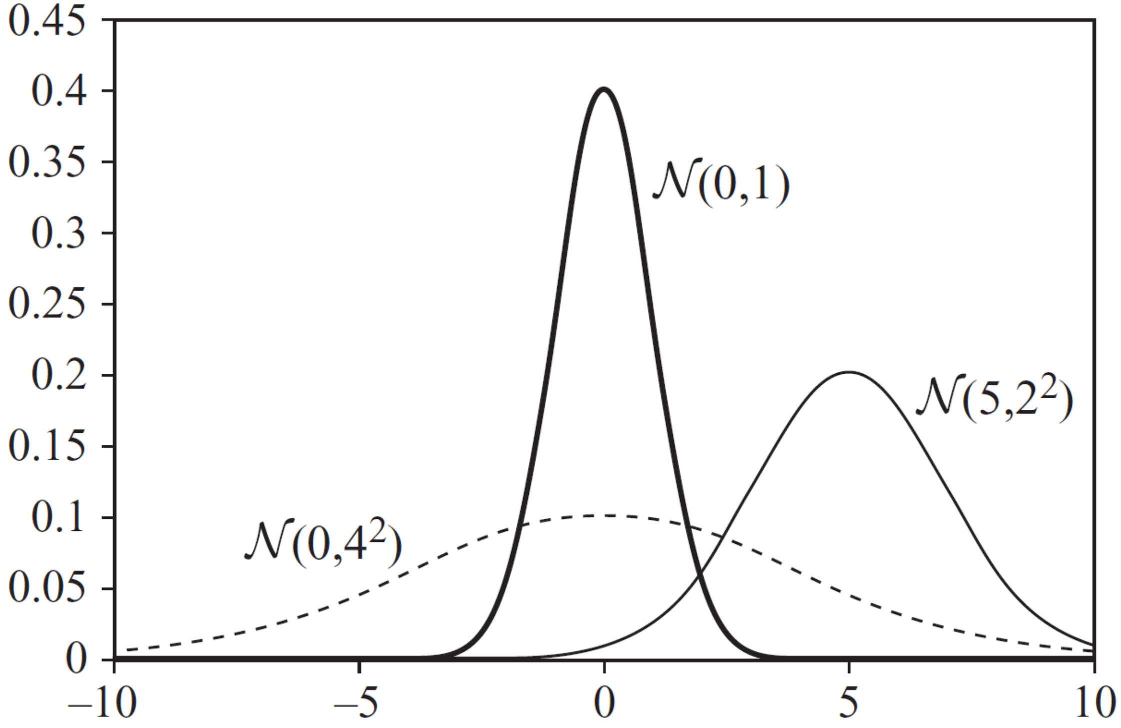
https://twitter.com/bilgicm

Motivation for this chapter

- So far, we focused primarily on discrete distributions
- The independence statements made by the structures, both Bayesian network structures and Markov network structures, apply regardless of whether the variables are discrete or real-valued
- In this chapter, we'll focus on one example
  - Gaussian network models

Univariate Gaussian Distribution

$$p(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



Multivariate Gaussian Distribution

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

Two-Variable Example

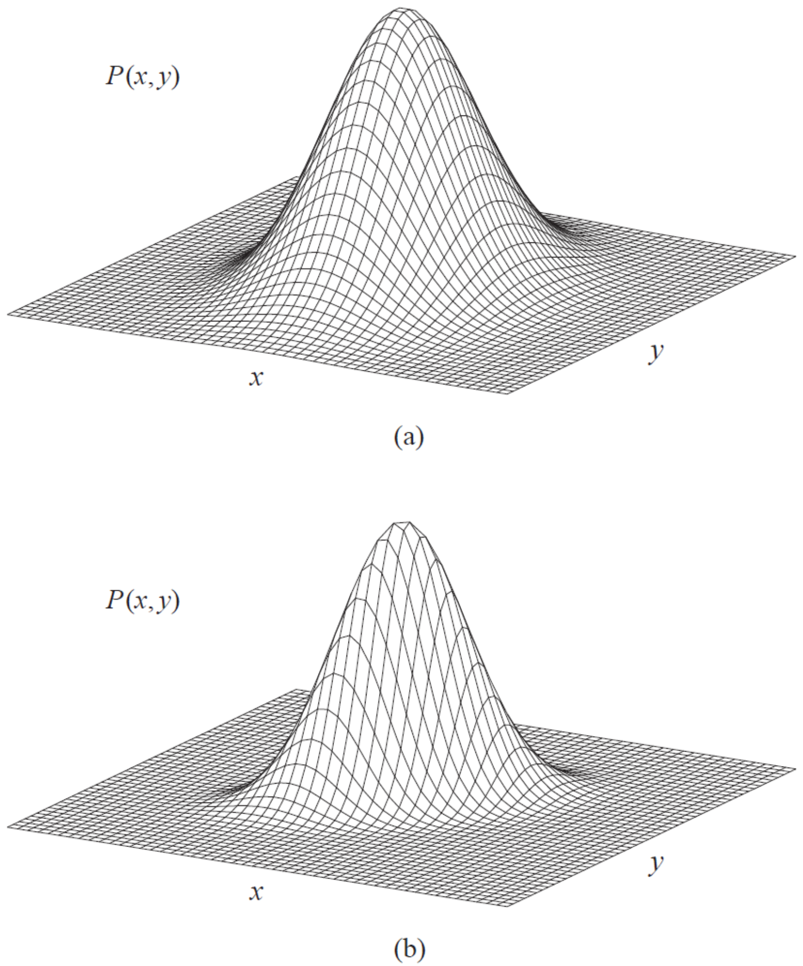


Figure 7.1 Gaussians over two variables X and Y. (a) X and Y uncorrelated. (b) X and Y correlated.

Three-Variable Example

$$\boldsymbol{\mu} = \begin{pmatrix} 1 \\ -3 \\ 4 \end{pmatrix} \qquad \Sigma = \begin{pmatrix} 4 & 2 & -2 \\ 2 & 5 & -5 \\ -2 & -5 & 8 \end{pmatrix}$$

Information Matrix J

$$
\begin{aligned}
-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) &= -\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T J(\boldsymbol{x}-\boldsymbol{\mu}) \\
&= -\frac{1}{2}\left[\boldsymbol{x}^T J\boldsymbol{x} - 2\boldsymbol{x}^T J\boldsymbol{\mu} + \boldsymbol{\mu}^T J\boldsymbol{\mu}\right]
\end{aligned}
$$

Theorem: $J_{ik} = 0$ if and only if X_i and X_k are independent given the rest of the variables.

INFERENCE

- Marginalization
  - Compute $p(\boldsymbol{X_s})$ for $\boldsymbol{X_s} \subseteq \boldsymbol{X}$
  - Use the covariance form
- Conditioning
  - Compute $p(\boldsymbol{X} \mid \boldsymbol{Z} = \boldsymbol{z})$
  - Use the information form

Three-Variable Example

$$\boldsymbol{\mu} = \begin{pmatrix} 1 \\ -3 \\ 4 \end{pmatrix} \qquad \Sigma = \begin{pmatrix} 4 & 2 & -2 \\ 2 & 5 & -5 \\ -2 & -5 & 8 \end{pmatrix}$$

$$J = \begin{pmatrix} 0.3125 & -0.125 & 0 \\ -0.125 & 0.5833 & 0.3333 \\ 0 & 0.3333 & 0.3333 \end{pmatrix}$$

Gaussian Bayesian Network

- Assume Y is a linear Gaussian of its parents $X_1, X_2, \ldots, X_k$

$$p(Y \mid \boldsymbol{x}) = \mathcal{N}\left(\beta_0 + \boldsymbol{\beta}^T \boldsymbol{x}; \sigma^2\right)$$

- If $X_1, X_2, \ldots, X_k$ are jointly Gaussian, then
  - $p(Y)$ is a Gaussian
  - $p(X, Y)$ is a Gaussian

Example

- Structure
  - $X_1 \rightarrow X_2 \rightarrow X_3$
- Parameters
  - $p(X_1) = N(1; 4)$
  - $p(X_2 \mid X_1) = N(-3.5 + 0.5X_1; 4)$
  - $p(X_3 \mid X_2) = N(1 - X_2; 3)$
- What is $p(X_1, X_2, X_3)$?

$$\boldsymbol{\mu} = \begin{pmatrix} 1 \\ -3 \\ 4 \end{pmatrix} \qquad \Sigma = \begin{pmatrix} 4 & 2 & -2 \\ 2 & 5 & -5 \\ -2 & -5 & 8 \end{pmatrix} \quad J = \begin{pmatrix} 0.3125 & -0.125 & 0 \\ -0.125 & 0.5833 & 0.3333 \\ 0 & 0.3333 & 0.3333 \end{pmatrix}$$