Chapter 8 Graphical Models

Yue Yu

8.1 Bayesian Networks

(8.15)

$$
\begin{aligned}
\mathbb{E}[x_i] &= \mathbb{E}\Bigg[\sum_{j\in \mathrm{pa}_i} w_{ij}\mathbb{E}[x_j] + b_i\Bigg] \\
&= \sum_{j\in \mathrm{pa}_i} w_{ij}\mathbb{E}[\mathbb{E}[x_j]] + \mathbb{E}[b_i] \\
&= \sum_{j\in \mathrm{pa}_i} w_{ij}\mathbb{E}[x_j] + b_i.
\end{aligned}
$$

(8.16)

$$
\begin{aligned}
\mathrm{var}[x_i, x_j] &= \mathbb{E}[(x_i - \mathbb{E}[x_i])(x_j - \mathbb{E}[x_j])] \\
&= \mathbb{E}\left[(x_i - \mathbb{E}[x_i])\left\{\sum_{k\in \mathrm{pa}_j} w_{jk}x_k + b_j + \sqrt{v_j}\epsilon_j - \left(\sum_{k\in \mathrm{pa}_i} w_{jk}\mathbb{E}[x_k] + b_j\right)\right\}\right] \\
&= \mathbb{E}\left[(x_i - \mathbb{E}[x_i])\left\{\sum_{k\in \mathrm{pa}_j} w_{jk}(x_k - \mathbb{E}[x_k]) + \sqrt{v_j}\epsilon_j\right\}\right] \\
&= \sum_{k\in \mathrm{pa}_j} w_{jk}\mathbb{E}[(x_i - \mathbb{E}[x_i])(x_k - \mathbb{E}[x_k])] + \mathbb{E}[(x_i - \mathbb{E}[x_i])\sqrt{v_j}\epsilon_j] \\
&= \sum_{k\in \mathrm{pa}_j} w_{jk}\mathbb{E}[(x_i - \mathbb{E}[x_i])(x_k - \mathbb{E}[x_k])] + \mathbb{E}\left[\left\{\sum_{l\in \mathrm{pa}_i} w_{il}(x_l - \mathbb{E}[x_l]) + \sqrt{v_i}\epsilon_i\right\}\sqrt{v_j}\epsilon_j\right] \\
&= \sum_{k\in \mathrm{pa}_j} w_{jk}\mathrm{cov}[x_i, x_k] + \sqrt{v_j}\,\mathbb{E}\left[\sum_{l\in \mathrm{pa}_i} w_{il}(x_l - \mathbb{E}[x_l])\right]\mathbb{E}[\epsilon_j] + \sqrt{v_i v_j}\,\mathbb{E}[\epsilon_i\epsilon_j] \\
&= \sum_{k\in \mathrm{pa}_j} w_{jk}\mathrm{cov}[x_i, x_k] + I_{ij}v_j
\end{aligned}
$$

where we assumed that ϵ is independent of x.

8.2 Conditional Independence

N/A

8.3 Markov Random Fields

N/A

8.4 Inference in Graphical Models

(8.52)

$$\begin{aligned}
p(x_n) &= \sum_{x_1}\cdots\sum_{x_{n-1}}\sum_{x_{n+1}}\cdots\sum_{x_N} p(\mathbf{x}) \\
&= \frac{1}{Z}\sum_{x_1}\cdots\sum_{x_{n-1}}\sum_{x_{n+1}}\cdots\sum_{x_N} \psi_{1,2}(x_1,x_2)\psi_{2,3}(x_2,x_3)\cdots\psi_{N-1,N}(x_{N-1},x_N) \\
&= \frac{1}{Z}\left[\sum_{x_{n-1}}\psi_{n-1,n}(x_{n-1},x_n)\cdots\left[\sum_{x_2}\psi_{2,3}(x_2,x_3)\left[\sum_{x_1}\psi_{1,2}(x_1,x_2)\right]\right]\right] \\
&\quad \left[\sum_{x_{n+1}}\psi_{n,n+1}(x_n,x_{n+1})\cdots\left[\sum_{x_N}\psi_{N-1,N}(x_{N-1},x_N)\right]\right].
\end{aligned}$$

(8.58)

$$\begin{aligned}
p(x_{n-1},x_n) &= \sum_{x_1}\cdots\sum_{x_{n-2}}\sum_{x_{n+1}}\cdots\sum_{x_N} p(\mathbf{x}) \\
&= \frac{1}{Z}\sum_{x_1}\cdots\sum_{x_{n-2}}\sum_{x_{n+1}}\cdots\sum_{x_N} \psi_{1,2}(x_1,x_2)\psi_{2,3}(x_2,x_3)\cdots\psi_{N-1,N}(x_{N-1},x_N) \\
&= \frac{1}{Z}\left[\sum_{x_{n-2}}\psi_{n-1,n-2}(x_{n-1},x_{n-2})\cdots\left[\sum_{x_2}\psi_{2,3}(x_2,x_3)\left[\sum_{x_1}\right]\right]\right]\psi_{n-1,n}(x_{n-1},x_n) \\
&\quad \left[\sum_{x_{n+1}}\psi_{n,n+1}(x_n,x_{n+1})\cdots\left[\sum_{x_N}\psi_{N-1,N}(x_{N-1},x_N)\right]\right] \\
&= \frac{1}{Z}\mu_\alpha(x_{n-1})\psi_{n-1,n}(x_{n-1},x_n)\mu_\beta(x_n).
\end{aligned}$$