Chapter 4
Graphical Models

4.1 Graphical Models

- GM for modelling
  - To describe how variables can interact, such as independence
  - Belief networks, Markov networks, chain graphs and influence diagrams
- GM for inference
  - To perform inference on a distribution
  - Factor graphs and junction trees

4.2 Markov Networks

- Definition 4.1 (Potential)
  - A potential $\phi(x)$ is a function with $\phi(x) \geq 0$
  - A joint potential $\phi(x_1, \ldots, x_n)$ is a function with $\phi(x_1, \ldots, x_n) \geq 0$

- A distribution is a potential satisfying normalization,

$$\sum_x \phi(x) = 1$$

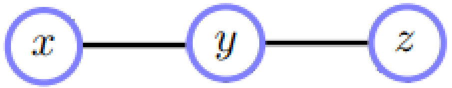
- Definition 4.2 (Markov Network).
  - For a set of variables $\mathcal{X} = \{x_1, \dots, x_n\}$
  - A Markov network is defined as a product of potentials on $\mathcal{X}_c \subseteq \mathcal{X}$:

$$p(x_1, \dots, x_n) = \frac{1}{Z} \prod_{c=1}^{C} \phi_c(\mathcal{X}_c)$$

  where

$$Z = \sum_{x_1, \dots, x_n} \prod_{c=1}^{C} \phi_c(\mathcal{X}_c)$$

  - Graphically this is represented by an undirected graph G with X_c, $c = 1, \dots, C$ being the maximal cliques of G.

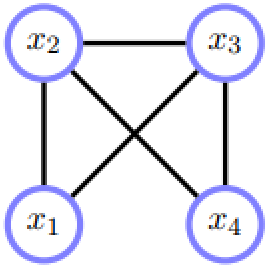


Maximal cliques: $\mathcal{X}_1 = \{x, y\}$, $\mathcal{X}_2 = \{y, z\}$

Potentials: $\phi(x, y)$, $\phi(y, z)$

Normalisation: $Z = \sum_{x,y,z} \phi(x, y)$, $\phi(y, z)$

Distribution: $p(x, y, z) = \phi(x, y)$, $\phi(y, z)/Z$



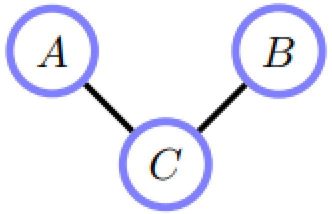
Maximal cliques: $\mathcal{X}_1 = \{x_1, x_2, x_3\}$, $\mathcal{X}_2 = \{x_2, x_3, x_4\}$

Potentials: $\phi(x_1, x_2, x_3)$, $\phi(x_2, x_3, x_4)$

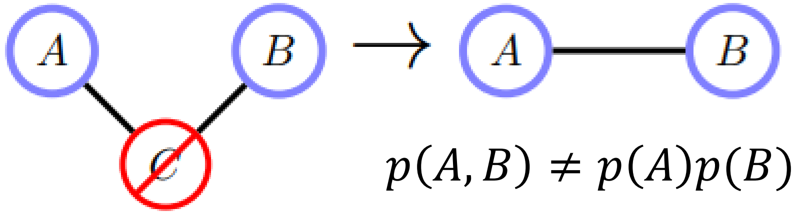
Normalisation: $Z = \sum_{x_1 x_2, x_3, x_4} \phi(x_1, x_2, x_3)$, $\phi(x_2, x_3, x_4)$

Distribution: $p(x_1, x_2, x_3, x_4) = \phi(x_1, x_2, x_3)$, $\phi(x_2, x_3, x_4)/Z$

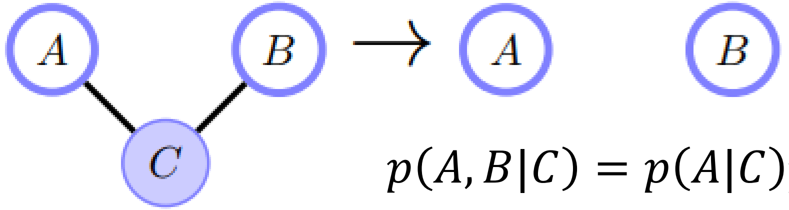
Definition 4.4 (Properties of Markov Networks).



$$p(A,B,C) = \phi_{AC}(A,C)\phi_{BC}(B,C)/Z$$



$$p(A,B) \neq p(A)p(B)$$



$$p(A,B|C) = p(A|C)p(B|C)$$

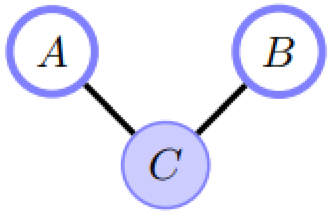
$$\begin{aligned}
p(A,B|C) &= \frac{p(A,B,C)}{p(C)} \\
&= \frac{p(A,B,C)}{\sum_{A',B'} p(A',B',C)} \\
&= \frac{\frac{1}{Z}\phi(A,C)\phi(B,C)}{\sum_{A',B'}\frac{1}{Z}\phi(A',C)\phi(B',C)} \\
&= \frac{\phi(A,C)}{\sum_{A'}\phi(A',C)}\frac{\phi(B,C)}{\sum_{B'}\phi(B',C)}
\end{aligned}$$

$$\begin{aligned}
p(A|C) &= \frac{p(A,C)}{p(C)} \\
&= \frac{\sum_{B'} p(A,B',C)}{\sum_{A',B'} p(A',B',C)} \\
&= \frac{\sum_{B'}\frac{1}{Z}\phi(A,C)\phi(B',C)}{\sum_{A',B'}\frac{1}{Z}\phi(A',C)\phi(B',C)} \\
&= \frac{\phi(A,C)}{\sum_{A'}\phi(A',C)}
\end{aligned}$$

$$p(A,B|C) = p(A|C)p(B|C)$$

- 4.2.1 Markov properties
  - properties of Markov networks, such as independence

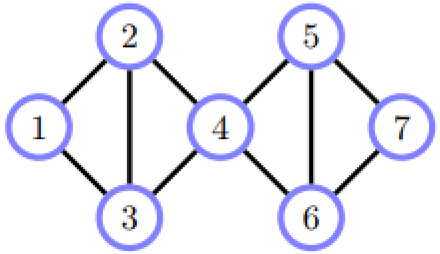
- Definition 4.5 (Separation)
  - A subset $\mathcal{S}$ separates a subset $\mathcal{A}$ from a subset $\mathcal{B}$ if every path from any member of $\mathcal{A}$ to any member of $\mathcal{B}$ passes through $\mathcal{S}$



C separates A from B

- Definition 4.6 (Global Markov Property)
  - If $\mathcal{S}$ separates $\mathcal{A}$ from $\mathcal{B}$, then $\mathcal{A}$ and $\mathcal{B}$ are independent given $\mathcal{S}$, i.e.

$$\mathcal{A} \perp\!\!\!\perp \mathcal{B} \mid \mathcal{S}$$



$1 \amalg 7 \mid 4$

Computational description

$$p(1,7|4) \propto \sum_{2,3,5,6} p(1,2,3,4,5,6,7)$$
$$= \frac{1}{Z} \sum_{2,3,5,6} \phi(1,2,3)\phi(2,3,4)\phi(4,5,6)\phi(5,6,7)$$
$$= \frac{1}{Z} \sum_{2,3} \phi(1,2,3)\phi(2,3,4) \sum_{5,6} \phi(4,5,6)\phi(5,6,7)$$
$$p(1|4) \propto \sum_{2,3,5,6,7} p(1,2,3,4,5,6,7)$$
$$= \frac{1}{Z} \sum_{2,3,5,6,7} \phi(1,2,3)\phi(2,3,4)\phi(4,5,6)\phi(5,6,7)$$
$$p(7|4) \propto \frac{1}{Z} \sum_{2,3,5,6,7} \phi(1,2,3)\phi(2,3,4)\phi(4,5,6)\phi(5,6,7)$$

using

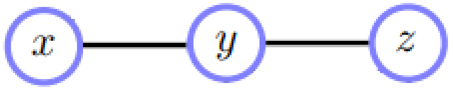
$$p(4) = \frac{1}{Z} \sum_{1,2,3} \phi(1,2,3)\phi(2,3,4) \sum_{5,6,7} \phi(4,5,6)\phi(5,6,7)$$

we obtain

$$p(1,7|4) = p(1|4)p(7|4)$$

- 4.2.2 Markov random fields
- Definition 4.7 (Markov Random Field)
  - A distribution is an MRF with respect to an undirected graph G if
    $$p(x_i \mid x_{\setminus i}) = p(x_i \mid ne(x_i))$$
    where $ne(x_i)$ are the neighbouring variables of variable x_i

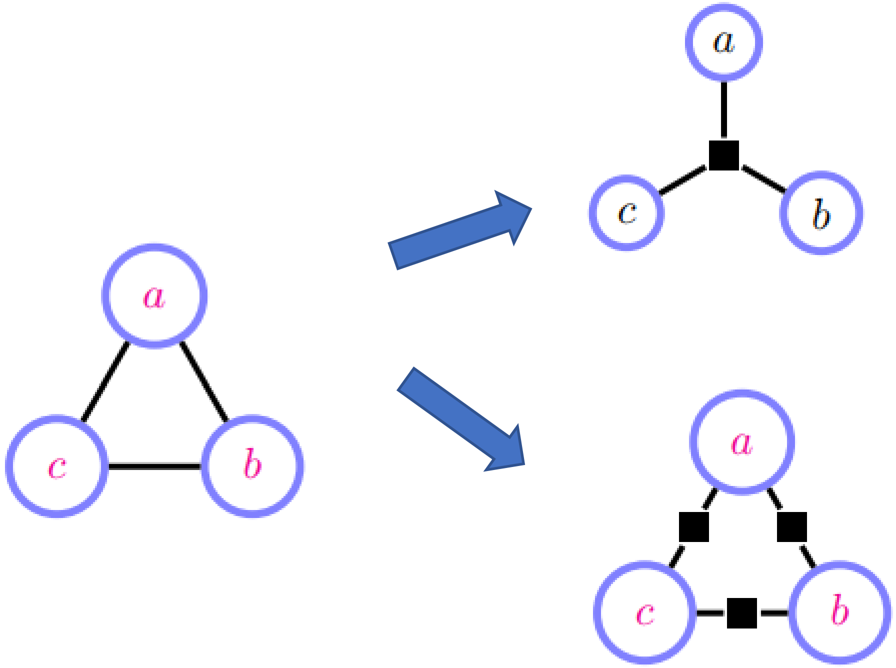
- A Markov network is a MRF
  - Eg.



$$p(x|y,z) = p(x|y)$$

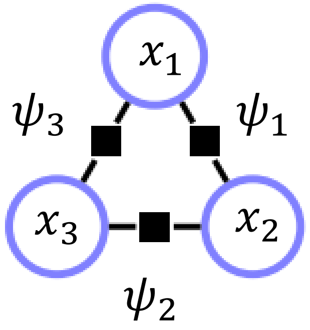
4.3 Chain Graphical Models

4.4 Factor Graphs



$$p(a,b,c) = \frac{1}{Z}\phi(a,b,c)$$

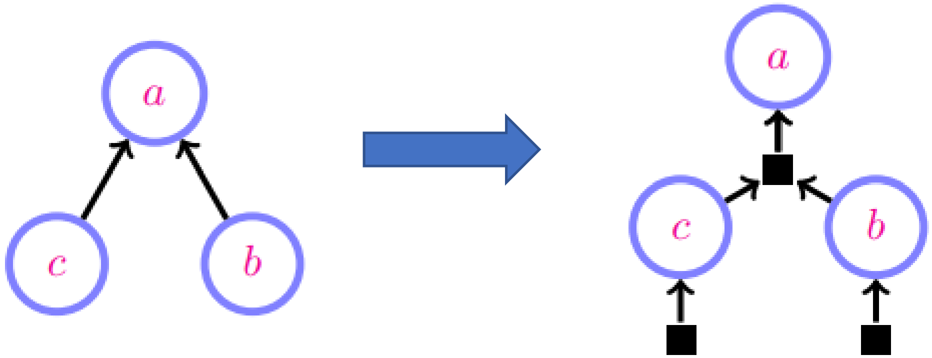
$$p(a,b,c) = \frac{1}{Z}\phi(a,b)\phi(b,c)\phi(c,a)$$



- Definition 4.10 (Factor Graph)
  - The FG has a node (a square) for each factor ψ_i and a variable node (a circle) for each variable x_j
  - For each $x_j \in \mathcal{X}_i$ an undirected link is made between factor ψ_i and variable x_j.
  - When used to represent a distribution

$$p(x_1, \dots, x_n) = \frac{1}{\mathrm{Z}} \prod_i \psi_i(\mathcal{X}_i)$$

    a normalisation constant $Z = \sum_{\mathcal{X}} \prod_i \psi_i(\mathcal{X}_i)$ is assumed. Here $\mathcal{X}$ represents all variables in the distribution.



$$p(a,b,c) = p(c|a,b)p(a)p(b)$$
$$= \frac{1}{Z}\phi(a,b,c)\phi(a)\phi(b)$$