Probabilistic Graphical Models Notes

Section 1: Background

Def 1 (Conditional Independence) *Given RVs X, Y, Z we say that $X \perp Y \mid Z$ iff $P(X, Y|Z) = P(X \mid Z)P(Y \mid Z)$*

Useful properties:

a) Symmetry: $X \perp Y \mid Z \Rightarrow Y \perp X \mid Z$

b) Decomposition: $X \perp Y, W \mid Z \Rightarrow \begin{cases} X \perp Y \mid Z \\ X \perp W \mid Y \end{cases}$

c) Weak Union: $X \perp Y, W \mid Z \Rightarrow \begin{cases} X \perp Y \mid Z, W \\ X \perp W \mid Z, Y \end{cases}$

d) Contraction: $\begin{cases} X \perp Y \mid Z, W \\ X \perp W \mid Z \end{cases} \Rightarrow X \perp Y, W \mid Z$

An ordering $I : v \mapsto \{1, \cdots, n\}$ is **Topological** iff $j \in \pi_i \Rightarrow I(j) < I(i)$ *i.e. parents always come before their children* If G is a DAG, then $\exists$ a Topological Ordering on G A **Forest** is a Graph s.t. each node has at most one parent A **Tree** is a Forest if it is connected

Section 2: Directed Graphical Models

Def 2 (Factorization Property of DAGs) *Given DAG $G = (V, E)$*

$$\mathcal{L}(G) = \left\{ p \text{ is a dist over } x_v : \exists \text{ factors } f_i \text{ s.t. } p(x_v) = \prod_{i=1}^n f_i(x_i; x_{\pi_i}) \text{ and } f_i \text{ satisfies: } \begin{matrix} f_i > 0 \\ \sum_{x_i} f_i(x_i; x_{\pi_i}) = 1 \\ f_i : Dom(x_i)^2 \mapsto [0, 1] \end{matrix} \right\}$$

Prop 1 (Leaf Plucking Property) *if n is a leaf of G, $p(x_v) \in \mathcal{L}(G - \{n\})$*

Proof. $P(x_v) = p(x_{1:n-1}, x_n) = f_n(x_n; x_{\pi_n}) \prod_{i \neq n} f_i(x_i, x_{\pi_i})$. Next marginalizing out x_n and using that it is a leaf, we have:

$$p(x_{1:n-1}) = \underbrace{\sum_{x_n} f_n(x_n; x_{\pi_n})}_{\text{sums to 1}} \underbrace{\prod_{i=1}^{n-1} f_i(x_i; x_{\pi_i})}_{\text{Does not contain } x_n} = \prod_{i=1}^{n-1} f_i(x_i; x_{\pi_i})$$

So $p(x_v) \in \mathcal{L}(G - \{n\})$ as needed. □

Prop 2 (Factors are Conditional PMFs) *Let $p \in \mathcal{L}(G)$ and $\{f_j\}$ be a factorization, then $\forall i, P(x_i|x_{\pi_i}) = f_i(x_i; x_{\pi_i})$*

Proof. WLOG let $\{1, \cdots, n\}$ be a Topological Ordering and use Theorem 1 to get that $p(x_{1:i}) \in \mathcal{L}(G - \{i + 1, \cdots, n\})$, and so $p(x_{1:i}) = \underbrace{\prod_{j=1}^{i-1} f_j(x_j; x_{\pi_j})}_{\text{Call this } g(x_{1:i-1})} f_i(x_i; x_{\pi_i})$. We partition $\{1 : i\}$ as $\{i\} \cup \pi_i \cup A$ and get that:

$$p(x_i \mid x_{\pi_i}) = \frac{\sum_{x_A} f_i(x_i; x_{\pi_i}) g(x_{1:i-1})}{\sum_{x_A} \sum_{x'_i} f_i(x'_i; x'_{\pi_i}) g(x_{1:i-1})} = \frac{f_i(x_i; x_{\pi_i}) \sum_{x_A} g(x_{1:i-1})}{\sum_{x'_i} f_i(x'_i; x'_{\pi_i}) \sum_{x_A} g(x_{1:i-1})} = f_i(x_i; x_{\pi_i})$$

□

Note: adding edges adds more distributions i.e. $E \subseteq E'$ and $G' = (V, E')$ then $\mathcal{L}(G) \subseteq \mathcal{L}(G')$

Prop 3 $p \in \mathcal{L}(G) \iff x_i \perp x_{nd(i)} \mid \pi_i$

Proof. $(\Rightarrow)$ $(\Leftarrow)$ □

References