Def. [Factorization (F)]

we say $P(x)$ factorizes according to G if

$$P(x) = \frac{1}{Z} \prod_{c \in \mathcal{C}} f_c(X_c)$$

where $\mathcal{C}$ is the set of maximal cliques in G.

Claim. (F) $\Rightarrow$ (G).

proof. for any $A-B-C$ separated in G

there is no clique that includes $a \in A$ and $c \in C$, hence

$$P(x) = \frac{1}{Z} F_1(X_A, X_B) \cdot F_2(X_B, X_C)$$

by HW1. this implies

$$X_A \perp\!\!\!\perp X_C \mid X_B.$$

Theorem. [Hammerseley-Clifford theorem]

If $P(x) > 0$ for all x, then (G) $\Rightarrow$ (F).

Implication: for $P(x) > 0$, (F) $\Leftrightarrow$ (G) $\Leftrightarrow$ (L) $\Leftrightarrow$ (P)

Counter example: If $P(x) \not> 0$, then (G) $\not\Rightarrow$ (F) (HW1).

Proof of Hammersley-Clifford theorem:

suppose $P(x) > 0$ & $X_A \perp\!\!\!\perp X_C \mid X_B$ for all A-B-C

then $P(x) \propto \prod_{c \in \mathcal{C}} f_c(x_c)$.

Def. pseudo factors

$$\tilde{f}_S(x_S) \triangleq \prod_{U \subseteq S} P(x_U, 0_{V \setminus U})^{(-1)^{|S \setminus U|}}$$

where we assumed w.l.o.g that $0 \in \mathcal{X}$ and we used $P(x) > 0$.

ex) singleton pseudo factors.

$\binom{n}{1}$ pseudo factors $\longrightarrow$ $\tilde{f}_i(x_i) = \dfrac{P(x_i, 0_{rest})}{P(0_i, 0_{rest})}$ $\quad x_{V \setminus \{i\}} = 0$ $\quad x_i = 0$

pairwise pseudo factors (for all i, j not necessarily an edge)

$\binom{n}{2}$ pseudo factors $\longrightarrow$ $\tilde{f}_{ij}(x_i, x_j) = \dfrac{P(x_i, x_j, 0_{rest}) \cdot P(0_i, 0_j, 0_{rest})}{P(x_i, 0_j, 0_{rest}) \cdot P(0_i, x_j, 0_{rest})}$

Claim 1. $\tilde{f}_S(x_S) =$ constant unless S is a clique.

claim 2. for any $P(x)$, $P(x) = P(0) \cdot \prod_{S \subseteq V} \tilde{f}_S(x_S)$

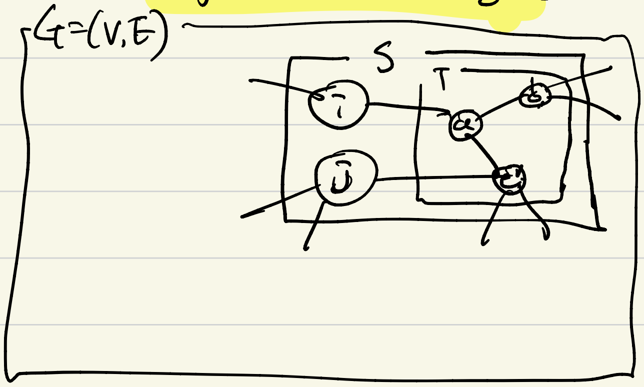
it follows from claims 1 & 2 that

$$P(x) \propto \prod_{c \in \mathcal{C}} \tilde{f}_c(x_c)$$

we are left to show claims 1 & 2

proof of claim 1. $\tilde{f}_S(x_S) = \text{Const}$ for S not clique & (G)

If S is not clique $\longrightarrow$ $\exists\, i,j$ not connected by an edge.



$S = T \cup \{i,j\}$, note that

subsets of S = subsets of T $= \left[\begin{array}{c} \emptyset \\ \{a,b,c\} \\ \{a\}\{b\}\{c\} \\ \{a,b\}\{b,c\}\{a,c\} \end{array} \right]$

$+$

$\{i\} \cup$ subsets of T

$+$

$\{j\} \cup$ subsets of T

$+$

$\{i,j\} \cup$ subsets of T.

$$\tilde{f}_S(x_S) \triangleq \prod_{U \subseteq S} P(x_U, O_{rest})^{(-1)^{|S \setminus U|}}$$

$$= \prod_{U \subseteq T} \underbrace{\left(\frac{P(x_U, x_i, x_j, O_{rest})\, P(x_U, O_i, O_j, O_{rest})}{P(x_U, x_i, O_j, O_{rest})\, P(x_U, O_i, x_j, O_{rest})} \right)}_{\text{constant}}{}^{(-1)^{|T \setminus U|}}$$

$$\frac{P(x_U, x_i, x_j, O_{rest})}{P(x_U, x_i, O_j, O_{rest})} \underset{(G):\ x_i \perp\!\!\!\perp x_j \mid x_{rest}}{=} \frac{\cancel{P(x_i \mid x_U, O_{rest})} \cdot P(x_U, x_j, O_{rest})}{\cancel{P(x_i \mid x_U, O_{rest})} \cdot P(x_U, O_j, O_{rest})}$$

$$= \frac{P(O_i \mid x_U, O_{rest})}{P(O_i \mid x_U, O_{rest})} \cdot \frac{P(x_U, x_j, O_{rest})}{P(x_U, O_j, O_{rest})}$$

$$= \frac{P(x_U, O_i, x_j, O_{rest})}{P(x_U, O_i, O_j, O_{rest})}$$

$\square$

Proof of claim 2. $P(x) \propto \prod_{S \subseteq V} \tilde{f}_S(X_S)$ if $P(x) > 0$

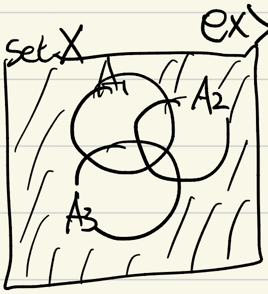
Recall that $\tilde{f}_S(X_S) = \prod_{U \subseteq S} P(X_U, O_{rest})^{(-1)^{|S \setminus U|}}$

Lemma [Möbius inversion lemma]. introduced in number theory 1832

for any $g, h : \{\text{subset of } V\} \to \mathbb{R}$, we have

$$g(S) = \sum_{U \subseteq S} h(U), \quad \text{for all } S \subseteq V$$

$$\Longleftrightarrow \quad h(S) = \sum_{U \subseteq S} (-1)^{|S \setminus U|} \cdot g(U), \quad \text{for all } S \subseteq V$$



ex) (inclusion-exclusion lemma) given sets A_1, A_2, A_3, $V = \{1,2,3\}$

$$|A_1^c \cap A_2^c \cap A_3^c| = |X| - |A_1| - |A_2| - |A_3| + |A_1 \cap A_2| + |A_2 \cap A_3| + |A_1 \cap A_3| - |A_1 \cap A_2 \cap A_3|$$

$g(1) = |A_1^c \cap A_2 \cap A_3|$, $h(S)$: # of elements in every set A_i, for $i \notin S$

$$g(\{1,2,3\}) = |A_1^c \cap A_2^c \cap A_3^c| = \sum_{U \subseteq \{1,2,3\}} (-1)^{|S \setminus U|} \cdot f(U) = \underset{U=S}{|X|} - |A_3| - |A_2| - |A_1| \underset{U=\{1,2\}}{} + \underset{U=\{1\}}{|A_2 \cap A_3|} \cdots - \underset{U=\emptyset}{|A_1 \cap A_2 \cap A_3|}$$

let $g(S) \triangleq \log P(X_S, O_{rest})$ (using positivity)

$$h(S) \triangleq \sum_{U \subseteq S} (-1)^{|S \setminus U|} \log P(X_U, O_{rest})$$

$$= \log \tilde{f}_U(X_U)$$

from Möbius inversion lemma,

$$g(V) \triangleq \log P(X_V) = \sum_{U \subseteq V} h(U) \triangleq \sum_{U \subseteq V} \log \tilde{f}_U(X_U)$$

$$P(X_V) = \prod_{U \subseteq V} \tilde{f}_U(X_U) \qquad \square$$

Recap:

Directed Graphical Models

Factorization

$$P(x) = \prod_{i=1}^{N} P^{(i)}(X_i \mid X_{\pi_i})$$

$\Updownarrow$

Global Markov Property

A-B-C d-separated

$X_A \perp\!\!\!\perp X_C \mid X_B$

$\Updownarrow$

Local Markov Property

$X_i \perp\!\!\!\perp X_{nd_i \setminus \pi_i} \mid X_{\pi_i}$

$\Updownarrow$

Ordered Markov Property

$X_i \perp\!\!\!\perp X_{pr_i \setminus \pi_i} \mid X_{\pi_i}$

Undirected Graphical Models

Factorization

$P(x)>0 \Updownarrow$ $P(x) = \frac{1}{Z} \prod_{c \in \mathcal{C}} f_c(X_c)$

$\Downarrow$

Global Markov Property

A-B-C Separated

$P(x)>0 \Updownarrow$ $X_A \perp\!\!\!\perp X_C \mid X_B$

$\Downarrow$

Local Markov Property

$P(x)>0 \Updownarrow$ $X_i \perp\!\!\!\perp X_{rest} \mid X_{\partial i}$

$\Downarrow$

pairwise Markov Property

$X_i \perp\!\!\!\perp X_j \mid X_{rest}, \forall (i,j) \notin E$

Claim: $\exists\, P(x)$ s.t. $\nexists\, G$ (directed or undirected) s.t. $I(G) = I(P)$

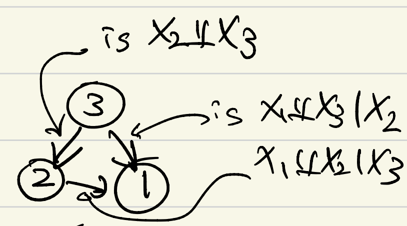
Def. G is an I-map of $P(x)$ if $I(G) \underset{=}{\subseteq} I(P)$.

G is a P-map = Perfect Map

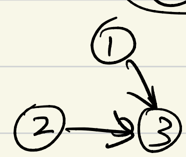
Def. G is an minimal I-map if removing any edge make it no longer I-map.

For BNs, you can construct an minimal I-map by fixing an ordering and removing edges.

ex) $V = (X_1, X_2, X_3)$, $X_1 \perp\!\!\!\perp X_2$, ordering $(3,2,1)$



ordering $(1,2,3)$



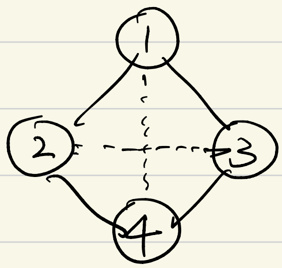
For MRFs, you remove edges (at arbitrary order) that can be removed.

Claim. If $P(x)>0$, minimal I-map is unique.

proof> by (F)$\Leftrightarrow$(P), removing an edge as per (F) does not create any independences not present in (F).

$V=\{1,2,3\}$

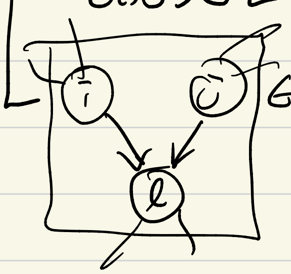
$X_1 \perp\!\!\!\perp X_3 \mid (X_2, X_4)$, $X_2 \perp\!\!\!\perp X_4 \mid (X_1, X_3)$



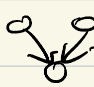
Definition. [Moralization]

For a DAG $G=(V,E)$, <u>moralization</u> of G is an undirected graph $G'=(V,E')$ where

$(i,j) \in E \Longrightarrow (i,j) \in E'$



$\in E \Rightarrow (i,j) \in E'$

complete all  -structure

Claim A DAG G's <u>moralization</u> does not add any edges,

$\Longleftrightarrow$ $\exists$ undirected G' s.t. $I(G) = I(G')$

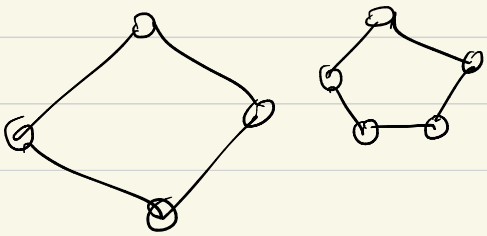
Definition. [Chordal Graph]

An undirected graph $G=(V,E)$ is chordal if every loop of size ≥ 4 has a chord

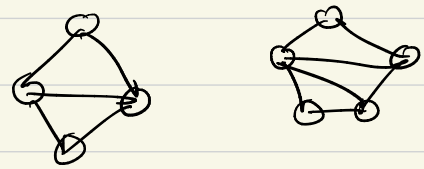
$\Downarrow$

an edge connecting 2 non-consecutive nodes in a loop

chordless



chordal



claim. undirected G is chordal

$\iff \exists$ DAG G' s.t. $I(G) = I(G')$.