Def. [Factorization (F)]

we say $P(x)$ factorizes accordly 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 $\underset{\substack{\uparrow \\ \mathcal{X}^n}}{x}$, 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).

(P, G) s.t. $P(x) > 0$, P satisfy (G) on G.

$\Longrightarrow$ find factorization s.t.

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

Proof of Hammersley-Clifford theorem: Given $P(x)$, $G=(V,E)$

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

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

Def. pseudo factors

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

$x_{S \setminus U} = \vec{0}$

$0 \in \mathcal{X}$

where we use $P(x)>0$.

ex> Singleton Pseudo factors.

all rest zeros.

$\binom{n}{1}$ singletons $\rightarrow$

$$\tilde{f}_i(x_i) = \frac{P(x_i, 0_{rest})}{P(0_i, 0_{rest})}$$

$\leftarrow U=\{i\}$

$\leftarrow U=\emptyset$

$S=\{i\}$ $\quad$ $x_i=0$ $\quad$ $x_{V \setminus \{i\}}=0$

cond $\rightarrow (0,1,0,2,3)$

$x_1\ x_2\ x_3\ x_4\ x_5$

ex> Pairwise Pseudo factors

$\binom{n}{2}$ pairwise $\longrightarrow$

$$\tilde{f}_{ij}(x_i, x_j) = \frac{P(x_i, x_j, 0_{rest}) \cdot P(0_i, 0_j, 0_{rest})}{P(x_i, 0_j, 0_{rest}) \cdot P(0_i, x_j, 0_{rest})}$$

$S=\{i,j\}$

claim 1. If $P(x)>0$, S is not a clique, and (G)

$$\Rightarrow \tilde{f}_S(x_S) = \text{Constant function.}$$

does not depend on x_S

claim 2. If $P(x)>0$

$$\Rightarrow P(x) = P(0) \cdot \prod_{S \subseteq V} \tilde{f}_S(x_S)$$

by claim 1 $\Longrightarrow = \frac{1}{Z} \prod_{c \in \mathcal{C}} \tilde{f}_c(x_c)$ $\square$.

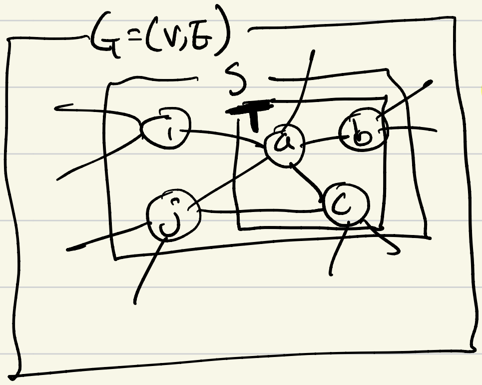
$P(x)>0$,

proof of claim 1. If S is not clique & $P(x)$ satisfy (G)

$$\tilde{f}_S(x_S) \triangleq \prod_{U\subseteq S} P(X_U, O_{rest})^{(-1)^{|S\setminus U|}} \quad \text{is constant.}$$

$$= \prod_{U\subseteq T}\left(\boxed{\frac{P(X_U, X_i, X_j, O_{rest})\cdot P(X_U, O_i, O_j, O_{rest})}{P(X_U, X_i, O_j, O_{rest})\cdot P(X_U, O_i, X_j, O_{rest})}}\right)^{(-1)^{|T\setminus U|}}$$

$= \text{Constant}.$



S is not clique

has i, j not an edge.

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

subsets of S

= subsets of T
+
$\{i\}\cup$ subsets of T
+
$\{j\}\cup$ subsets of T
+
$\{i,j\}\cup$ subsets of T

= $\left\{\begin{matrix}\emptyset, \{a\}, \{b\}, \{c\} \\ \{a,b\}, \{b,c\}, \{c,a\} \\ \{a,b,c\}\end{matrix}\right\}$

$$\frac{A\cdot P(X_U, X_i, X_j, O_{rest})}{B\cdot P(X_U, X_i, O_j, O_{rest})} = \frac{P(X_i \mid \cancel{X_U, O_{rest}})\cdot P(X_U, X_j, O_{rest})}{P(X_i \mid \cancel{X_U, O_{rest}})\cdot P(X_U, O_j, O_{rest})}$$

i, j not edge

(G): $X_i \perp X_j \mid X_{rest}$.

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

$$= \frac{P(X_U, O_i, X_j, O_{rest})\cdot C}{P(X_U, O_i, O_j, O_{rest})\cdot D}$$

$$\tilde{f}_S(x_S) = \prod_{U\subseteq T}\left(\underbrace{\frac{A\cdot D}{B\cdot C}}_{=1}\right)^{(-1)^{|T\setminus S|}} = \text{Constant}.$$

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 $g, h : \{\text{subset of } V\} \to \mathbb{R}$, the following are equivalent
>
> $$g(S) = \sum_{U \subseteq S} h(U), \quad \text{for all } S \subseteq V$$
>
> $$\Longleftrightarrow \quad h(S) = \sum_{U \subseteq S} (-1)^{|S \setminus U|} \cdot g(U), \quad \text{for all } S \subseteq V.$$

for $P(x) > 0$.

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

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

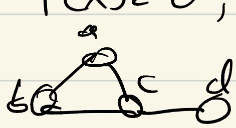
$$= \log \underbrace{\prod_{U \subseteq S} P(X_S, O_{rest})^{(-1)^{|S \setminus U|}}}_{\tilde{f}_S(X_S)} = \log \tilde{f}_S(X_S)$$

möbius inversion lemma.

$$g(V) = \log(P(X_V)) \stackrel{\triangle}{=} \sum_{S \subseteq V} h(S) = \sum_{S \subseteq V} \log \tilde{f}_S(X_S) = \log \prod_{S \subseteq V} \tilde{f}_S(X_S)$$

$$P(x) = \prod_{\substack{S \subseteq V \\ \varnothing}} \tilde{f}_S(X_S) = \qquad \square.$$

If $P(x) > 0$, then $(G) \Longrightarrow (F)$.



proof $\to$ $P(x) \propto \boxed{f(x_a, x_b) \times f(x_b, x_c) \times f(x_a, x_c) \times f_{abc}(x_a, x_b, x_c)}\, f(x_c, x_d)$

show $\to$ $P(x) \propto f_{abc}(x_a, x_b, x_c) \cdot f(x_c, x_d)$

Recap:

Directed Graphical Models

Factorization

$P(x) = \prod_{i=1} P(x_i | x_{\pi_i})$

$\Updownarrow$

Global M.P

A–B–C d-separated

$X_A \perp\!\!\!\perp X_C | X_B$

$\Updownarrow$

Local M.P

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

$\Updownarrow$

Ordered M.P

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

Undirected Graphical Models

Factorization

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

$\Downarrow$ ($\Uparrow$ $P(x)>0$)

Global M.P

A–B–C separated

$X_A \perp\!\!\!\perp X_C | X_B$

$\Downarrow$ ($\Uparrow$ $P(x)>0$)

Local M.P

$X_i \perp\!\!\!\perp X_{rest} | X_{\partial i}$

$\Downarrow$ ($\Uparrow$ $P(x)>0$)

Pairwise M.P

$X_i \perp\!\!\!\perp X_j | X_{rest}$, if ij not in E

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

Def: G is <u>I-map</u> of P(x) if $I(G) \subseteq I(P)$

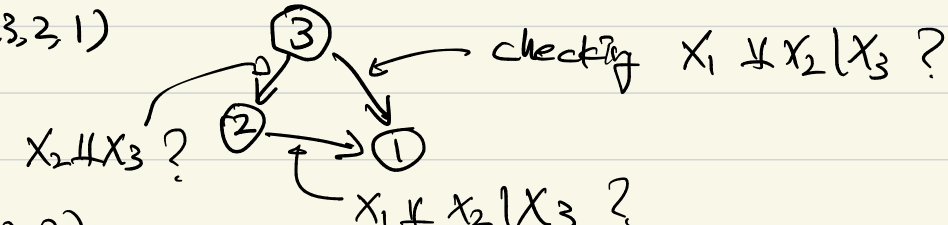
G is a P-map = Perfect Map

Def: G is <u>minimal I-map</u> if removing an edge causes it to be no longer I-map.

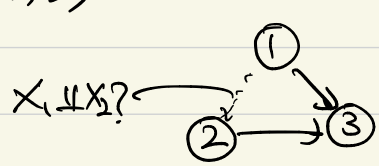
For BNs, you construct a minimal I-map by choose ordering, and remove edges.

ex) $V = \{1,2,3\}$, $P(X)$, $I(P) = \{X_1 \perp\!\!\!\perp X_2\}$.

① Ordering (3, 2, 1)



② ordering (1, 2, 3)

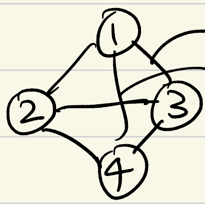


For MRF. remove any edge that can be removed in any order.

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

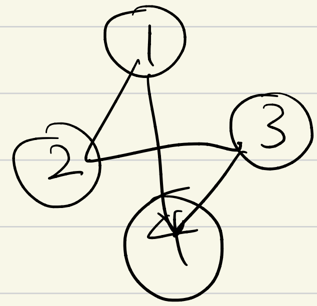
$V = \{1,2,3,4\}$ $\quad I(P) = \{ X_1 \perp\!\!\!\perp X_3 \mid (X_2, X_4)$

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



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

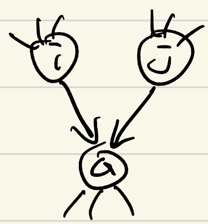
$X_1 \perp\!\!\!\perp X_4 \mid (X_2, X_3)$? NO

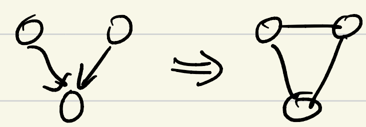


? DAG $\longrightarrow$ Undirected.

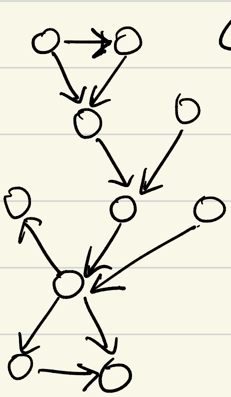
Def. [Moralization]

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

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



G

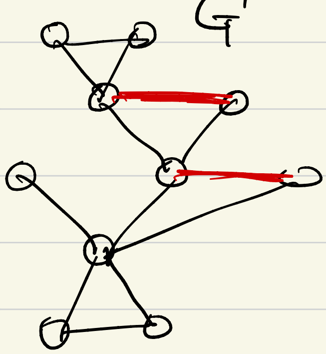


claim: $I(\text{DAG } G) \supseteq I(\text{moral } G')$

claim: If moralization does not add any edges,

$\Rightarrow I(G) = I(G')$

G'



? Undirected $G \Longrightarrow$ DAG

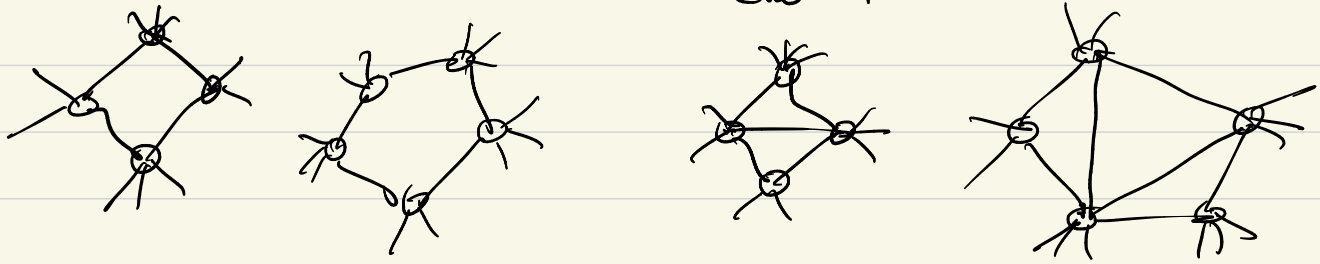
Definition [Chordal Graph]

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

loop = path that ends where it begins.

chord $\equiv$ an edge connecting 2 non-consecutive nodes in a loop.

chordless | chordal



claim. undirected G is chordal.
$\Longleftrightarrow \exists$ DAG G' s.t $I(G)=I(G')$.