A directed graphical model on $G=(V,E)$ $\Rightarrow$ a DAG
is a family of distributions that factorizes as

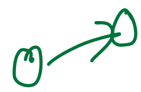
$$\Pr[x_1, \dots x_n] \triangleq \prod_{i=1}^{n} f_i(x_i, x_{\pi(i)}) \qquad (\exists f_i)$$

$$\text{s.t.} \quad \sum_{x_i \in \mathcal{X}} f_i(x_i, x_{\pi(i)}) = 1$$

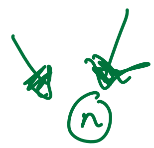
Claim: $f_i(x_i, \pi_{x_i}) = P(x_i, x_{\pi(i)})$ $\forall i \in V$ $\quad P(\cdot|\cdot) \Rightarrow$ Probability Mass fn

WLOG

Suppose $G=(V,E)$ has topological ordering $\{1, \dots, n\}$ s.t. $i \to j$, $(i,j) \in E \Rightarrow i <_o j$



$$(1) \quad P(x_1^n) := P(x_1, \dots, x_n) = \prod_i f_i(x_i, x_{\pi(i)})$$



$$(2) \quad P(x_1^{n-1}) = P(x_1, \dots, x_{n-1}]$$

$$= \sum_{x_n \in \mathcal{X}} P(x_1, \dots, x_n)$$

$$= \sum_{x_n \in \mathcal{X}} P(x_1^n)$$

$$= \sum_{x_n \in \mathcal{X}} \prod_{i=1}^{n} f_i(x_i, x_{\pi(i)}) \qquad [\text{By } (1)]$$

$$= \sum_{x_n \in \mathcal{X}} f_n(x_n, x_{\pi(n)}) \cdot \prod_{i=1}^{n-1} f_i(x_i, x_{\pi(i)})$$

$$= \underbrace{\left(\sum_{x_n \in \mathcal{X}} f_n(x_n, x_{\pi(n)})\right)}_{=1 \text{ by def}} \cdot \prod_{i=1}^{n-1} f_i(x_i, x_{\pi(i)}) \qquad (\text{Since } x_{\pi(i)} \not\ni x_n \text{ By T/O})$$

$$= \prod_{i=1}^{n-1} f_i(x_i, x_{\pi(i)})$$

$$f_n(x_n, x_{\pi(n)}) = \frac{P(x_1^n)}{P(x_1^{n-1})}$$ (By def of DGM)

$$= P(x_n \mid x_1^{n-1})$$ (Def of CPD)

$$= P(x_n \mid x_{\pi(n)})$$ [Since $f_n(x_n, x_{\pi(n)})$ Only depends on $x_n, x_{\pi(n)}$]

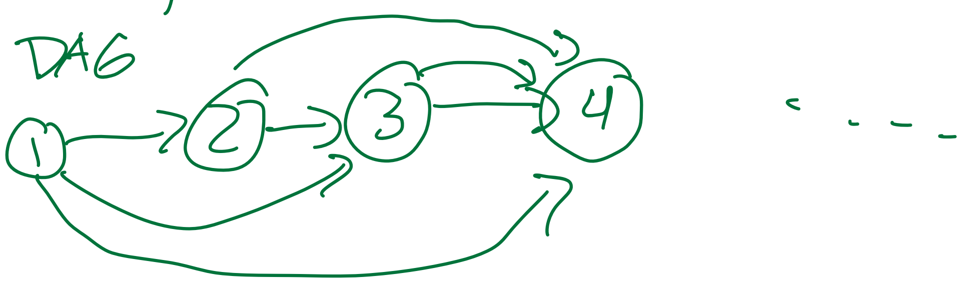
For the remaining nodes, proof by induction on remaining graph (removing n). $\square$

Directed GMs encode Bayes' Rule / aka the Chain Rule

For any $P(X)$ & ordering $(1, \dots, n)$

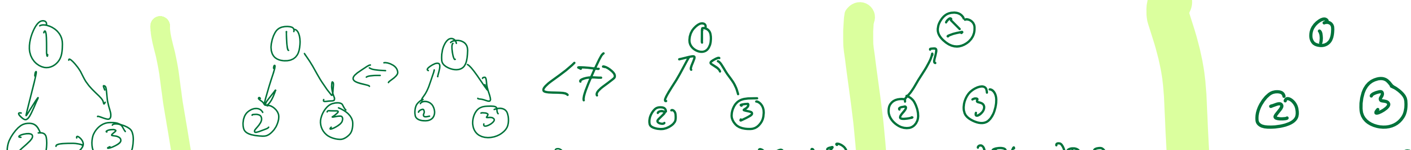
$$P(x) = P(x_1) P(x_2 \mid x_1) P(x_3 \mid x_2, x_1) \dots P(x_n \mid x_1^{n-1})$$

Any $P(x)$ is representable by the complete DAG



$(i, j) \in E \quad \forall i < j$

What dependencies does a DAG encode?



$P(x_1)P(x_2|x_1)P(x_3|x_1,x_2)$ | $P(x_1)P(x_2|x_1)P(x_3|x_1)$ $\Leftrightarrow$ $P(x_2)P(x_1|x_2)P(x_3|x_1)$ $\not\Leftrightarrow$ $P(x_2)P(x_3)P(x_1|x_2,x_3)$ | $P(x_2)P(x_1|x_2)P(x_3)$ | $P(x_1)P(x_2)P(x_3)$

$\emptyset$ | $x_3 \perp\!\!\!\perp x_2 | x_1$ | $x_3 \perp\!\!\!\perp x_2$ | $x_3 \perp\!\!\!\perp (x_1, x_2)$ | All mutually $\perp\!\!\!\perp$

dense complex not any conditional Indep $\longleftrightarrow$ sparse, simple structured

---

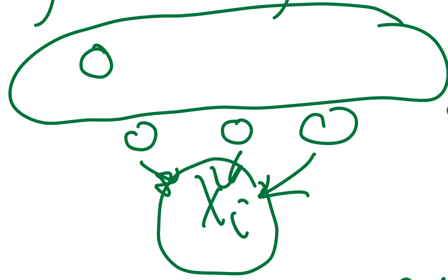
Markov Properties of Directed graphical models

(i) Directed Ordered Markov Property

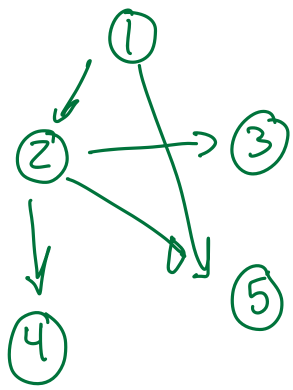
Given a DAG $G=(V,E)$ p $(1,\dots,n)$

$$X_i \perp\!\!\!\perp X_{Pr_i \setminus \pi(i)} \mid X_{\pi(i)}$$

$Pr(i)$ = nodes before i



Pf: Chain rule $\square$



Ordering $(1,2,3,4,5)$

$x_3 \perp\!\!\!\perp x_1 | x_2$

$x_4 \perp\!\!\!\perp x_1, x_3 | x_2$

$x_5 \perp\!\!\!\perp x_1, x_3 \mid x_1, x_2$

Ordering $(1,2,4,5,3)$

$x_4 \perp\!\!\!\perp x_1 \mid x_2$

$x_3 \perp\!\!\!\perp x_1, x_4, x_5 | x_2$

$\vdots$

---

(2) Directed Local Markov Property

$$X_i \perp\!\!\!\perp X_{nd(i) \setminus \pi(i)} \mid \pi(i)$$

$nd(i)$ = nondescendents of i = set of nodes j where $\nexists$ directed path $i, \dots, j$

x_1 $\emptyset$

x_2 $\emptyset$

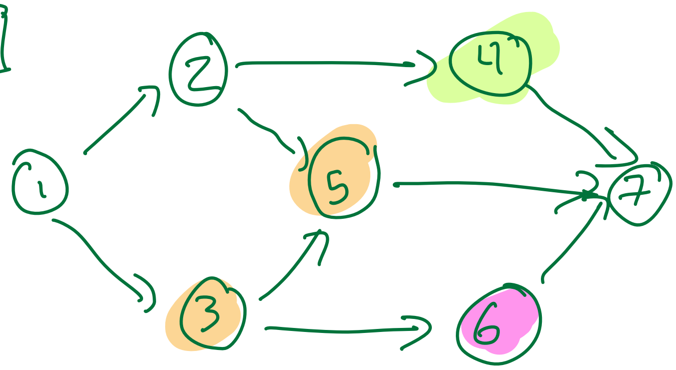
$x_3 \perp\!\!\!\perp x_1, x_4, x_5 \mid x_2$

Q: How do you determine all CI given a G?

∃ CI Test(A, B, C, G) → {0,1} where CI(A,B,C,G) = 1 iff $A \perp\!\!\!\perp C \mid B$ in G?

A: Yes, the Bayes Ball Alg.

$\boxed{G}$



A = {4}

B = {3, 5}

C = {6}

is $A \perp\!\!\!\perp C \mid B$?

③ Directed Global Markov Property

For any $A, B, C \subseteq V$ $\quad X_A \perp\!\!\!\perp X_C \mid X_B$ if A is d-separated by B from C

A is d-separated by B from C if ∀ undirected, simple (nonrepeating) paths $a_1 \ldots c$ ($a_1 \in A$, $c \in C$) the path is blocked by B.

Path $a_1 \ldots c$ is blocked by B if ∃ 3 nodes

$\boxed{\text{Intuition}}$

a) .... ① → ② → ③ .... c $\quad$ P $2 \in B$ $\quad x_1 \perp\!\!\!\perp x_3 \mid x_2$

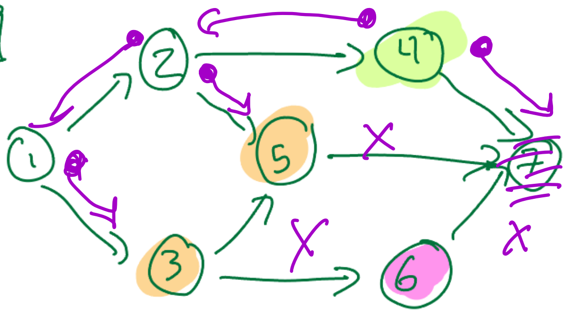
a) .... ① ← ② ← ③ .... c $\quad$ P $2 \in B$ $\quad x_1 \perp\!\!\!\perp x_3 \mid x_2$

a) .... ① ← ② → ③ .... c $\quad$ P $2 \in B$ $\quad x_1 \perp\!\!\!\perp x_3 \mid x_2$

a) .... ① → ② ← ③ .... c $\quad$ P $2 \notin B$ $\quad x_1 \perp\!\!\!\perp x_3$

Bayes' Ball identifies all CI in a graph

$\boxed{G}$



$A = \{4\}$

$B = \{3, 5\}$

$C = \{6\}$

7 acts as a block $\notin B$

Thm: All of the following are equivalent

(a) $P(x)$ factorizes according to $G = (V, E)$

(b) $P(x)$ satisfies the directed global Markov prop

(c) $P(x)$ satisfies the DAG local Markov prop

$\boxed{\text{Pf}}$

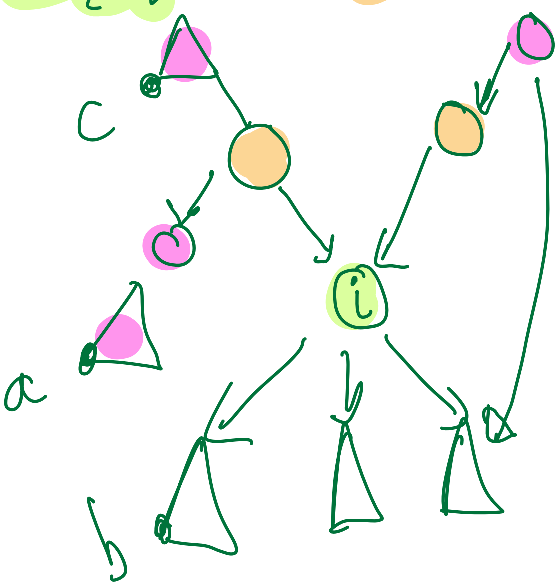
$b \Rightarrow c$

$A = \{i\}$ $\quad B = \pi_i$ $\quad C = \text{Nd}\, i \setminus \pi_i$

$c \perp\!\!\!\perp i \mid \pi_i$

$a \perp\!\!\!\perp i \mid \pi_i$

$b \not\perp\!\!\!\perp i \mid \pi_i$



descendents of i

$c \Rightarrow a$ any topological ordering $(1, \dots, n)$

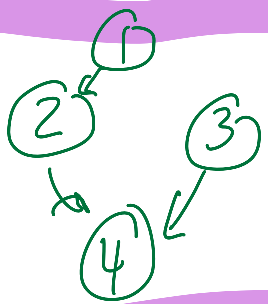
$$P(x) = \prod_i P(x_i \mid x_1^{i-1}) = \prod_i P(x_i \mid x_{\pi(i)})$$

chain

local MP

$a \Rightarrow b$ in a separate PDF online (by ind)

Let $I(G)$ = set of all $\perp\!\!\!\perp$ implied by $G=(V,E)$



$$I(G) = \{ X_3 \perp\!\!\!\perp X_1, X_2$$
$$X_4 \perp\!\!\!\perp X_1 \mid X_2, X_3 \}$$

Let $I(D)$ be all $\perp\!\!\!\perp$ implied by factorization of P

Consider $Z_1 \perp Z_2 \sim \text{Bern}(1/3)$

XOR

$X_1 = Z_1 \quad X_2 = Z_1 \quad X_3 = Z_2, \; X_4 = X_2 \oplus X_3$

$$I(P) = \{ X_3 \perp\!\!\!\perp X_1, X_2$$
$$X_1 \perp\!\!\!\perp X_3, X_4 \mid X_2$$
$$X_2 \perp\!\!\!\perp X_3, X_4 \mid X_1 \}$$

Remarks ① $\exists P(X)$ s.t. no DAG perfectly captures $\perp\!\!\!\perp$ (ie $I(G) \neq I(P)$)

[Equality of RV's is a weird edge case]

$X_1 = X_2$ $\qquad$ $X_1 \perp\!\!\!\perp X_3 \mid X_2$

$X_3 = X_1 + Z$ $\qquad$ $X_2 \not\!\perp\!\!\!\perp X_3 \mid X_1$

(2) $\forall G, \exists P$ s.t. $I(G) = I(P)$

(3) $\exists G_1 \neq G_2$ s.t. $I(G_1) = I(G_2)$

$\circ \rightarrow \circ \rightarrow \circ \qquad \circ \rightarrow \circ \leftarrow \circ$

<u>Def</u> G is a <u>perfect map</u> for P if $I(P) = I(G)$

<u>Def</u> G is an <u>I-map</u> for $P(X)$ if $I(G) \subseteq I(P)$

<u>Def</u> G is a minimal I-map for $P(X)$ if

1) Remove an edge in $G \Rightarrow$
G no longer an I-map

(4) Min I-MAP isn't unique

- different ordering of vars
- even w = ordering,

$$\begin{tikzcd} 1 \arrow[r] & 2 \arrow[r] & 3 \end{tikzcd} \qquad \begin{tikzcd} 1 \arrow[r] \arrow[rr, bend left] & 2 & 3 \end{tikzcd}$$

$x_1 = x_2$

$x_3 = x_1 + z$