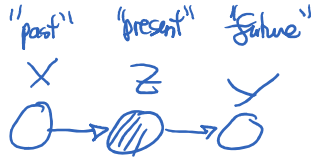


today: • DGM  
• UGM

conditional indep. in DGM

Basic 3 nodes graphs

1) Markov chain



here  $Y \perp\!\!\!\perp X \mid Z$

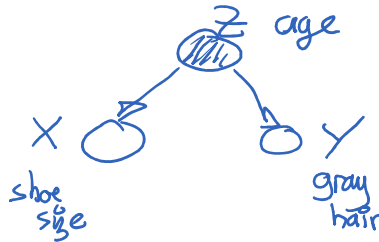
but  $X \not\perp\!\!\!\perp Y$  (for some  $P \in \mathcal{L}(G)$ )

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

(exercise)

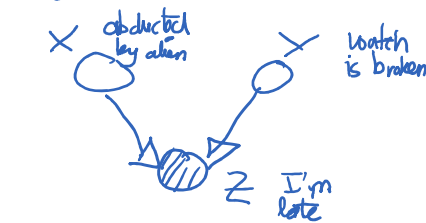
$$\Rightarrow p(z, y|x) = p(x|z)p(y|z)$$

2) latent cause (hidden variable)



same statements as above

3) explaining away / competing effect (v structure)



$X \perp\!\!\!\perp Y$

but  $X \not\perp\!\!\!\perp Y \mid Z$  (for some  $P$ )

non-monotonic property of conditioning &

$$p(\text{alien}) \text{ tiny}$$

$$p(\text{alien} \mid \text{late}) > p(\text{alien})$$

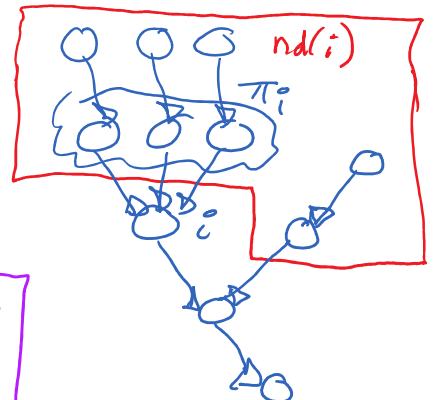
$$p(\text{alien} \mid \text{late}, \text{watch broken}) < p(\text{alien} \mid \text{late})$$

more cond. independence statements in a DGM

$$\text{let } \text{nd}(i) \triangleq \{j \neq i : \text{no path } i \rightarrow j\}$$

"non-descendants" of  $i$

prop:  $P \in \mathcal{L}(G) \Leftrightarrow X_i \perp\!\!\!\perp X_{\text{nd}(i)} \mid X_{\pi_i} \forall i$



(by decomposition:  $\Rightarrow X_i \perp\!\!\!\perp X_j \mid X_{\pi_i} \forall j \in nd(i)$ )

proof  $\Rightarrow$ ) Key point: let  $i$  be fixed

then  $\exists$  a top. ordering s.t.  $nd(i)$  are exactly before  $i$

i.e.  $(nd(i), i, \text{descendant}(i))$

pluck all these?

$$\Rightarrow p(x_i, x_{nd(i)}) = p(x_i | x_{\pi_i}) \prod_{j \in nd(i)} p(x_j | x_{\pi_j})$$

$$\frac{p(x_i | x_{nd(i)})}{p(x_i | x_{\pi_i}, x_{nd(i) \setminus \pi_i})} = \frac{p(x_i, x_{nd(i)})}{p(x_{nd(i)})} = \frac{p(x_i | x_{\pi_i}) \prod_{j \in nd(i)} p(x_j | x_{\pi_j})}{\left( \sum_{x_i} p(x_i | x_{\pi_i}) \right) \prod_{j \in nd(i)} p(x_j | x_{\pi_j})} = p(x_i | x_{\pi_i})$$

no  $x_i$ 's here

$\Leftarrow$ ) (suppose  $p$  satisfies all cond. indep statements)

let  $1:n$  be (WLOG) a top. sort of  $G$

then  $1:i-1 \subseteq nd(i) \forall i$  (by top sort property)

$\Rightarrow X_i \perp\!\!\!\perp X_{1:i-1} \mid X_{\pi_i}$  by decomposition

$$p(x_i) = \prod_{j=1}^i p(x_j | x_{1:j-1}) \quad (\text{by chain rule})$$

this includes  $\pi_i$

$$= \prod_{j=1}^i p(x_j | x_{\pi_j}) \quad (\text{by cond. indep.})$$

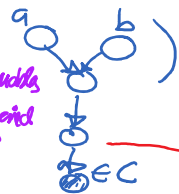
$$\Rightarrow p \in \mathcal{P}(G)$$

14640

other cond. indep statements?

chain from  $a$  to  $b$ : just any (undirected) path in graph from  $a$  to  $b$

d-separation: (intuition: do not want this situation)



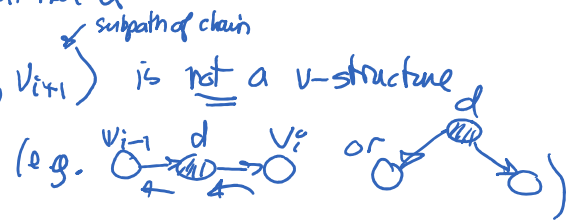
def: set  $A$  &  $B$  are said to be d-separated by  $C$  (in  $G$ )

iff all chain chains  $a \in A$  to  $b \in B$  are "blocked" given  $C$

where a chain  $a$  to  $b$  is "blocked" given  $C$  at node  $d$   
subpath of chain

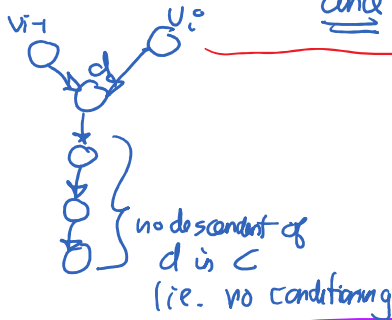
where a chain  $a$  to  $b$  is "blocked" given  $C$  at node  $d$

if a) either  $d \in C$  and  $(v_{i-1}, d, v_{i+1})$  is not a v-structure



b)  $d \notin C$  and  $(v_{i-1}, d, v_{i+1})$  is a v-structure

and no descendant of  $d$  is in  $C$



prop:

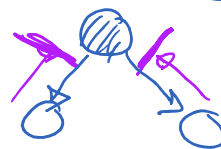
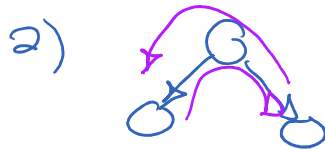
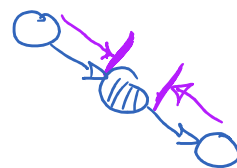
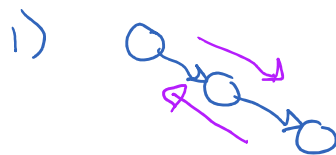
$$p \in \mathcal{P}(G) \Leftrightarrow X_A \perp\!\!\!\perp X_B \mid X_C \quad \forall A, B, C \subseteq V$$

s.t.  $A$  &  $B$  are d-separated given  $C$

"Bayes Ball" alg. "intuition" alg to check d-separation

rules for balls/chains being blocked

a) in d-separation

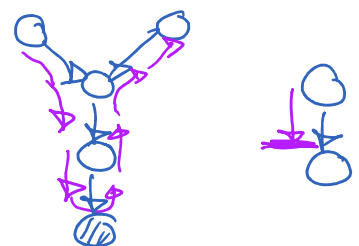


b)

3) v-structures



see Alg. 3.1  
in K&F book



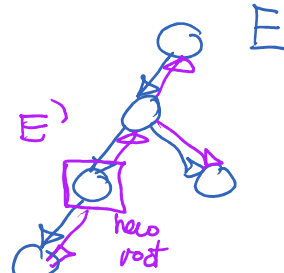
## Properties of DGM:

inclusion:  $E \subseteq E' \Rightarrow \mathcal{J}(G) \subseteq \mathcal{J}(G')$

reversal: if  $G$  is a directed tree (or a forest)  
 $\hookrightarrow |r_i| \leq 1$  i.e. there is no v-structure

let  $E'$  be another directed tree  
 (with some undirected edges) by  
 choosing a different root

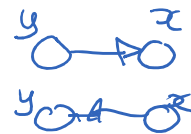
(reverse some edges, while keeping a  
 directed tree)



$$\Rightarrow \mathcal{J}(G) = \mathcal{J}(G')$$

(\*) why direction of edges  
 are not causal in DGM

rephrasing: all directed trees from an undirected tree  
 give same DGM



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

## marginalization:

- marginalizing a leaf node  $n$  gives us a smaller DGM

let  $S = \{q \text{ is a dist. on } x_{1:n-1} \text{ s.t. } q(x_{1:n-1}) = p(x_{1:n-1})\}$   
 for some  $p \in \mathcal{J}(G)$

then  $S = \mathcal{J}(G')$  where  $G'$  is  
 $G$  with leaf  $n$  plucked  
 (removed)

- not true for all marginalization

e.g.



$\hookrightarrow$  get a set of dist.  
 $\neq$  to any  $\mathcal{J}(G)$  for some  $G'$

i.e. marginalization is not a "closed" operation on DGM



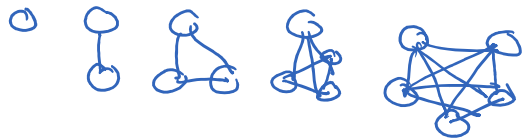
undirected GM (UGM) (aka. Markov random field or Markov network)

Undirected GM (UGM) (aka. Markov random field or Markov network)

let  $G = (V, E)$  be an undirected graph

let  $\mathcal{C}$  be the set of cliques of  $G$

clique = fully connected set of nodes  
 (if  $C \in \mathcal{C} \iff \forall i, j \in C, \exists i, j \in E$ )



UGM associated with  $G$

$$\mathcal{L}(G) \triangleq \left\{ p \text{ is a dist. over } X_V : p(x_V) = \frac{1}{Z} \prod_{C \in \mathcal{C}} \psi_C(x_C) \right\}$$

$$\psi_C : \Omega_{X_C} \rightarrow \mathbb{R}_+$$

for some "potentials"  $\psi_C$  s.t.  $\psi_C(x_C) > 0 \forall x_C$

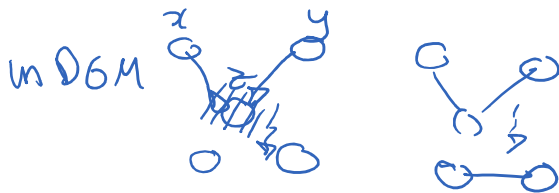
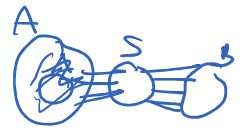
and  $Z \triangleq \sum_{x_V} \left( \prod_{C \in \mathcal{C}} \psi_C(x_C) \right)$   
 normalization constant  
 "partition function"

$$\sum_{x_1} \sum_{x_2} \sum_{x_3} \dots \sum_{x_n}$$

(we'll see next class that  $X_A \perp\!\!\!\perp X_B \mid X_S$  when  $S$  separates  $A$  from  $B$  in  $G$ )

you don't need  $\sum_{x_C} \psi_C(x_C) = 1$

eg.  $\psi_{\{1,2\}}(x_{1,2})$



$$p(x, y, z) = \frac{p(x, y, z)}{p(z)} = \frac{p(z|x, y) p(x) p(y)}{\sum_{x', y'} p(z|x', y') p(x') p(y')}$$

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