

today: continue DGM's properties

"leaf plucking" property fundamental property of DGM $p(x_r) = p(x_{1:n})$

let $p \in \mathcal{L}(G)$; let n be a leaf in G (i.e. n is not a parent of anything)

a) then $p(x_{1:n-1}) \in \mathcal{L}(G - \{n\})$

b) moreover, if $p(x_{1:n}) = \prod_{i=1}^n f_i(x_i | x_{\pi_i}) \Rightarrow p(x_{1:n-1}) = \prod_{i=1}^{n-1} f_i(x_i | x_{\pi_i})$
↳ means remove node n from G

proof: $p(x_n, x_{1:n-1}) = f_n(x_n | x_{\pi_n}) \prod_{j \neq n} f_j(x_j | x_{\pi_j})$
 $p(x_{1:n-1}) = \sum_{x_n} p(x_n, x_{1:n-1}) = \left(\sum_{x_n} f_n(x_n | x_{\pi_n}) \right) \left(\prod_{j \neq n} f_j(x_j | x_{\pi_j}) \right)$
↳ by definition ↳ no n in any π_j ↳ no x_n above

proposition: if $p \in \mathcal{L}(G)$

let $\{f_i\}$ be a factorization w.r. to G for p

then $\forall i, p(x_i | x_{\pi_i}) = f_i(x_i | x_{\pi_i}) \forall x_{\pi_i}$

i.e. factors are correct conditions
 (and are unique)

proof: WLOG, let $(1, \dots, n)$ be a topolog. sort of G (possible since G is a DAG)

let i be fixed; we want to prove that $p(x_i | x_{\pi_i}) = f_i(x_i | x_{\pi_i})$

• n is a leaf \rightarrow pluck it to get $p(x_{1:n-1}) = \prod_{j=1}^{n-1} f_j(x_j | x_{\pi_j})$

factorization for a DAG where x_{n-1} is now a leaf

• pluck $n-1$ as well?

• keep doing this [in reverse order $n-1, n-2, \dots, i+1$] to get

$$p(x_{1:i}) = f_i(x_i | x_{\pi_i}) \prod_{j: i_j < i} f_j(x_j | x_{\pi_j})$$

$\triangleq g(x_{1:i-1})$

ie. no x_i by top sort-property?

partition $1:i$ as $\{i\} \cup \pi_i \cup A$ [ie. $A \triangleq 1:i-1 \setminus \pi_i$]

$$p(x_{1:i}) = p(x_i, x_{\pi_i}, x_A)$$

$$p(x_i | x_{\pi_i}) = \sum_{x_A} p(x_i, x_{\pi_i}, x_A) = f_i(x_i | x_{\pi_i}) \left(\sum_{x_A} g(x_{1:i-1}) \right) = f_i(x_i | x_{\pi_i})$$

$\sum_{x_i} \sum_{x_A} p(x_i, x_{\pi_i}, x_A)$

~~$f_i(x_i | x_{\pi_i}) \cdot g(x_{1:i-1})$~~

x_i is not there

^(now) we can safely write

$$\mathcal{L}(G) = \left\{ p : p(x_v) = \prod_{i=1}^n p(x_i | x_{\pi_i}) \right\}$$

properties:

adding edges \Rightarrow more distributions

ie. $G = (V, E)$ and $G' = (V, E')$ with $E' \supseteq E$

then $\mathcal{L}(G) \subseteq \mathcal{L}(G')$

examples:

- $G_{a.b.}$
• $E = \emptyset$
(trivial graph)

$\Rightarrow \mathcal{L}(G)$ contains only fully independent distributions

ie. $p(x_v) = \prod_{i=1}^n p(x_i)$
 \uparrow
 $p(x_i | \emptyset)$

and $\mathcal{L}(G) \subseteq \mathcal{L}(G')$ for any G'

- complete digraph (DAG)
for all $i \neq j$ either $(i, j) \in E$
or $(j, i) \in E$



get wlog
 $p(x_v) = \prod_{i=1}^n p(x_i | x_{1:i-1})$
 (like chain rule)

\Rightarrow all dist are in $\mathcal{L}(G_{\text{complete}})$

removing edges impose more factorization constraints (and thus cond. indep. assumptions) on \mathcal{D}

conditional indep. in DGM

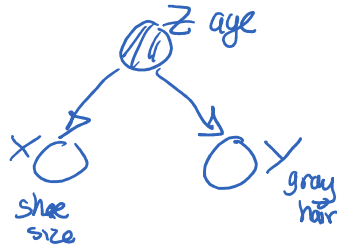
Basic 3 nodes graphs

1) Markov chain



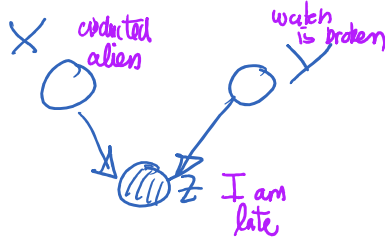
"future" "past" "present"
 here $Y \perp\!\!\!\perp X \mid Z$
 but $X \not\perp\!\!\!\perp Y$ (for some $p \in \mathcal{P}(G)$)

2) latent cause
 (hidden variable)



same statements as above
 i.e. $X \perp\!\!\!\perp Y \mid Z$
 $X \not\perp\!\!\!\perp Y$ (for some p)

3) explaining away competing effect
 (v-structure)



$X \perp\!\!\!\perp Y$
 but $X \not\perp\!\!\!\perp Y \mid Z$
 for some p

non-Markov property of conditioning is

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

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

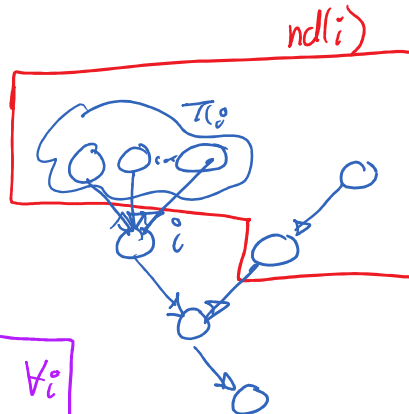
$p(\text{alien} \mid \text{late, broken watch}) < p(\text{alien} \mid \text{late})$

15h30

more cond. indep. statements in DGM

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

"non-descendants" of i



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

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

proof:

\Rightarrow) key point: let i be fixed

then \exists a tp. ordering st. $\text{nd}(i)$ are exactly before i

ie. $(\text{nd}(i), i, \text{descendent}(i))$
pluck all these?

\Rightarrow

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

$$p(x_i \mid x_{\text{nd}(i)}) = \frac{p(x_i, x_{\text{nd}(i)})}{p(x_{\text{nd}(i)})} = p(x_i \mid x_{\pi_i}) \frac{\prod_{j \in \text{nd}(i)} p(x_j \mid x_{\pi_j})}{\prod_{j \in \text{nd}(i)} p(x_j \mid x_{\pi_j})} = p(x_i \mid x_{\pi_i})$$

$\left(\sum_{x_i} p(x_i \mid x_{\pi_i}) \right) \prod_{j \in \text{nd}(i)} p(x_j \mid x_{\pi_j})$
 \downarrow
 1

*no x_i 's here
[because $j \in \text{nd}(i)$]
 $\Rightarrow i \in \pi_j$*

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

let $1:n$ be (w/LOG) a tp. sort of G

then $1:i-1 \subseteq \text{nd}(i) \forall i$ [by tp. sort. property]

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

$$p(X_V) = \prod_{i=1}^n p(x_i \mid x_{1:i-1}) \quad (\text{by chain rule})$$

\hookrightarrow this includes π_i

$$= \prod_{i=1}^n p(x_i \mid x_{\pi_i}) \quad (\text{by cond. indep.})$$

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

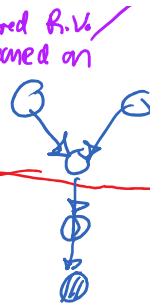
other cond. indep. statements?

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

d -separation: (intuition: is do not want to this situation \curvearrowright \curvearrowleft)
set of observed R.V. / conditioned on

d-separation:

(intuition: is do not want to ~~this situation~~ ^{set of observed R.V. / conditioned on})



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

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

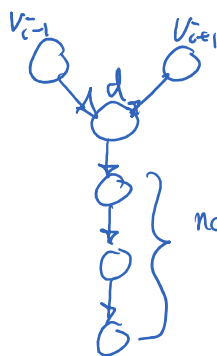
where a chain a to b is "blocked" at node d

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:

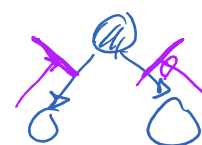
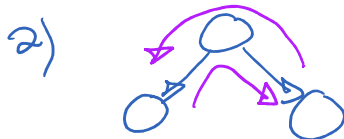
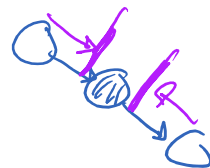
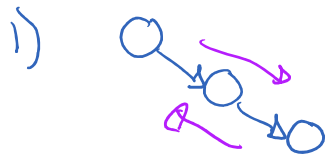
if $p \in \mathcal{L}(G) \Leftrightarrow X_A \perp\!\!\!\perp X_B \mid X_C \quad \forall A, B, C$

sl. A & B are d-separated given C

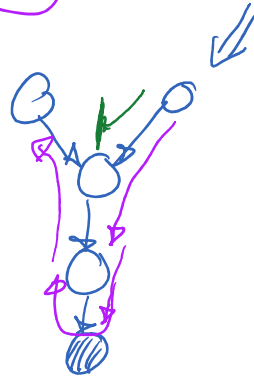
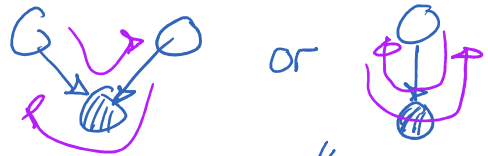
"Bayes Ball" alg: "intuitive" alg to check d-separation

rules for balls/chains being blocked

a) in d-separation



(b) 3) V-structure



see Alg. 3.1
in KIF book

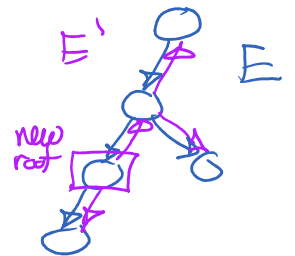
properties of DGM:

Inclusion: $E \subseteq E' \Rightarrow \mathcal{D}(G) \subseteq \mathcal{D}(G')$

Reversal: if G is a directed tree (or forest)

$\hookrightarrow |\pi_i| \leq 1$ i.e. there is no V-structure

If (E') be another directed tree (with same edges) by choosing a different root (reverse some edges, while keeping a directed tree)



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

⊛ this is why direction of edges are not causal in a DGM

[e.g. $p(x|y) = p(x|z)p(z|y) = p(y|z)p(x|z)$]



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

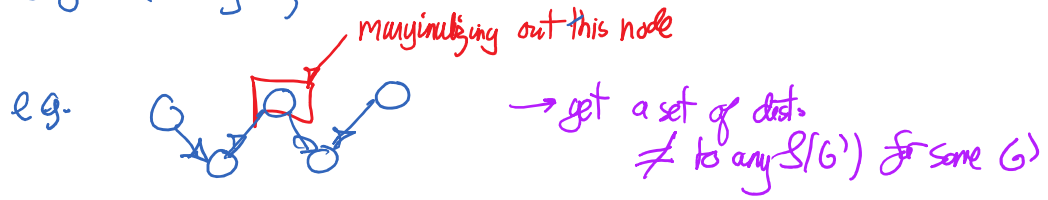
marginalization:

• marginalizing a leaf node n gives us a smaller DGM

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

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

• not true for all marginalization



i.e. marginalization is not a "closed operation" on $\mathcal{D}(M)$

